

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL



## **Uma abordagem à melhoria de resultados do produto automóvel**

Filipe Lopes Azinheira

**Mestrado em Matemática Aplicada à Economia e Gestão**

Trabalho de Projeto orientado por:  
Professora Doutora Teresa Alpuím  
Doutora Alexandra Mina

2018



## **Agradecimentos**

Gostaria de agradecer às minhas orientadoras, professora Teresa Alpuím e Alexandra Mina, por todo o apoio, disponibilidade e interesse durante a realização do projeto.

Gostaria também de agradecer à Companhia de seguros onde realizei este estágio pela oportunidade que me deram com este estágio e com a realização deste projeto.

À minha equipa de trabalho e ao responsável do ramo automóvel, o meu muito obrigado pela constante disponibilidade para me ajudar quer na realização deste projeto, quer durante o estágio.

Não posso deixar de agradecer a toda a minha família, especialmente aos meus pais, irmão e avós, pois sem eles não teria sido possível a realização deste percurso académico.

A todos os meus amigos e colegas que me acompanharam neste percurso.



## Resumo

Este relatório apresenta o trabalho desenvolvido no projeto em ambiente empresarial realizado no departamento de Atuariado *Pricing* Não Vida numa Companhia de seguros a operar em Portugal, com o objetivo de aplicar os conhecimentos adquiridos durante o mestrado em Matemática Aplicada à Economia e Gestão a um problema proposto por essa empresa, nomeadamente, a melhoria da gestão da carteira de seguro automóvel para o segmento individual.

Perante este objetivo, tornou-se fundamental conhecer em detalhe o passado recente do produto, quer a sua composição, quer alguns aspetos relacionados com a sua gestão. Assim, começou-se por estudar a evolução da carteira entre os anos 2014 e 2017 através da análise de diferentes variáveis (como por exemplo: zona geográfica, marca do veículo ou idade do condutor) utilizando indicadores chave específicos da atividade seguradora. A informação analisada provém de bases de dados gerais do ramo automóvel, contudo os dados apresentados no decorrer do projeto não são os reais, por questões de confidencialidade dos dados, pelo que foram multiplicados por um coeficiente. Utilizou-se os softwares SAS e Excel para tratamento e construção das tabelas resumo para as diferentes variáveis, que deram origem às análises efetuadas. Para se perceber melhor a evolução do produto, foi também feita uma análise de tendências para alguns dos indicadores chave e um estudo do impacto de algumas medidas tomadas pelo ramo automóvel da Companhia no período em estudo.

Recorrendo a métodos estatísticos, foram construídos dois programas em Excel para ajudar na gestão da carteira. O primeiro programa analisa todas as variáveis, de acordo com restrições previamente definidas, de modo a auxiliar na rápida identificação de segmentos onde é necessário agir, com o objetivo de melhorar a composição do portefólio e o resultado da carteira. O segundo programa cria grupos semelhantes entre si de modo a rapidamente identificar quais os segmentos com comportamento idêntico.

Por fim, refletiu-se sobre métodos alternativos de controlo do risco, nomeadamente sistemas avançados de apoio à condução (por exemplo travagem automática), que são sistemas que permitem ajudar na prevenção de sinistros, e telemáticas, que são sistemas em que o cliente paga em função do seu comportamento.

**Palavras-Chave:** Seguro Automóvel, Gestão da Carteira, Análise de Tendências, Análise de Variância

## Abstract

This report presents the work developed in the internship in the Pricing Non Life Actuarial department of the insurance company, with the objective of applying the knowledge acquired during the master's degree in Applied Mathematics to Economics and Management to a problem proposed by the company, namely, to find methods that can improve the management of the motor insurance portfolio for the non-fleet segment.

For this goal, it became fundamental to know in detail the recent past of the product, its composition and some aspects related to its management. Thus, the evolution of the portfolio between the years 2014 and 2017 was studied by analyzing different variables (such as geographical area, vehicle brand and driver age) using key indicators specific to the insurance activity. The information analyzed comes from motor non fleet databases. SAS and Excel softwares were used for the treatment and construction of the summary tables for the different variables that gave rise to the analyzes performed. In order to better understand the evolution, a trend analysis was also made for some of the key indicators and a study of the impact of some measures was taken by the company's motor non-fleet portfolio during the study period.

Using statistical methods, two Excel programs were built to help portfolio management. The first program analyzes all variables, according to previously defined restrictions, in order to help in the quick identification of segments where action is required, so as to improve the composition and the result of the portfolio. The second program creates similar groups in order to quickly identify which segments have identical behavior.

Finally, alternative risk control methods were explored, namely advanced driving support systems (e.g. automatic braking), which are systems that help to prevent accidents, as well as telematics, which are systems where the customer pays depending on his behavior.

**Key-Words:** Motor Insurance, Portfolio Management, Trend Analysis, Variance Analysis

# ÍNDICE

LISTA DE TABELAS	1
LISTA DE GRÁFICOS	2
LISTA DE FIGURAS	3
1. ATIVIDADE SEGURADORA	5
1.1. Breve história dos seguros	5
1.2. Principais Conceitos da Atividade Seguradora	6
1.3. Ramos da atividade seguradora	6
1.3.1. Ramo Vida	6
1.3.1.1. Seguros de Vida	6
1.3.1.2. Seguros de Nupcialidade ou Natalidade	7
1.3.1.3. Operações de Capitalização	7
1.3.1.4. Seguros Ligados a Fundos de Investimento	7
1.3.2. Ramo Não Vida	7
1.3.2.1. Ramo Automóvel	7
1.3.2.2. Ramo Transportes	7
1.3.2.3. Ramo Responsabilidade Civil	7
1.3.2.4. Ramo Patrimoniais	7
1.3.2.4.1. Seguro Multirriscos Habitação	8
1.3.2.4.2. Seguro Multirriscos Indústria e Comércio	8
1.3.2.5. Ramo Acidentes	8
1.3.2.5.1. Acidentes de Trabalho	8
1.3.2.5.2. Acidentes Pessoais	8
1.3.2.6. Ramo Saúde	8
2. SEGURO AUTOMÓVEL	9
2.1. Tipos de seguro	9
2.1.1. Responsabilidade Civil	9
2.1.2. Danos Próprios	9
2.1.3. Outros Danos	9
2.2. Fundo de garantia automóvel	10
2.3. Sistema de regularização de sinistros IDS	11
2.4. Bónus-Malus	11
3. MÉTODOS ESTATÍSTICOS UTILIZADOS	13
3.1. Intervalos de confiança	13
3.1.1. Intervalo de confiança para o valor médio, variância desconhecida	13
3.1.2. Intervalo de confiança para uma proporção	13

3.2.	Modelos de Regressão Linear	14
3.2.1.	Regressão Linear Simples	14
3.2.2.	Regressão Linear Múltipla	16
3.3.	Análise de Variância Simples	17
3.3.1.	Teste de Igualdade de Médias	17
3.3.2.	Análise de Variância	18
3.3.3.	Análise de variância como modelo de regressão	20
4.	ANÁLISE DA CARTEIRA AUTOMÓVEL DA COMPANHIA	24
4.1.	Indicadores utilizados	24
4.2.	Metodologia utilizada	25
4.3.	Primeiras conclusões dos resultados observados	27
4.4.	Análise de tendências e das medidas tomadas no passado	29
4.4.1.	Análise de Tendências	29
4.4.2.	Análise das medidas do passado	33
4.5.	Companhia em Análise vs. Mercado Segurador	38
5.	MÉTODOS DE GESTÃO DA CARTEIRA	40
5.1.	Tabela Resumo	40
5.2.	Programa para criar grupos	42
6.	SISTEMAS ALTERNATIVOS DE CONTROLO DE RISCO	46
6.1.	ADAS	46
6.2.	Telemáticas	48
7.	CONCLUSÃO	49
8.	REFERÊNCIAS	51
8.1.	Referências Bibliográficas	51
8.2.	Referências Sitográficas	51



## Lista de Tabelas

Tabela 4.1: Frequência de Pagamento.....	27
Tabela 4.2: Zona Geográfica do Agente.....	27
Tabela 4.3: Tipo de Cliente .....	28
Tabela 4.4: Tipo de Negócio .....	28
Tabela 4.5: Tipo de Combustível .....	28
Tabela 4.6: Tipo de Cliente e Tipo de Combustível.....	29
Tabela 4.7: Tabela Global com a Variação Anual .....	30
Tabela 4.8: Tipo de Pagamento Fracionado .....	34
Tabela 4.9: Impacto da redução da venda de apólices fracionadas .....	34
Tabela 4.10: Danos Próprios (DP) .....	35
Tabela 4.11: Impacto do aumento da tarifa de DP em 10% .....	35
Tabela 4.12: Sem Franquia.....	35
Tabela 4.13: Impacto da medida em apólices com DP sem franquias.....	36
Tabela 4.14: Franquia 2% .....	36
Tabela 4.15: Impacto da medida em apólices com DP com franquias 2% .....	36
Tabela 4.16: Apólices Saneadas.....	37
Tabela 4.17: Impacto das apólices saneadas .....	37
Tabela 4.18: Quota de Mercado por Distrito 2016.....	38
Tabela 5.1: Tabela para definir os limites utilizados.....	40
Tabela 5.2: Resumo dos resultados para o cálculo dos intervalos de confiança para custo médio .....	41
Tabela 5.3: Exemplo tabela resumo para 2017 .....	41
Tabela 5.4: Exemplo tabela resumo para os ‘piores’ segmentos em 2017 .....	42
Tabela 5.5: Tabela Final .....	45
Tabela 6.1: Impacto na frequência .....	48

## Lista de Gráficos

Gráfico 4.1: Exposição .....	30
Gráfico 4.2: Prémio Médio.....	31
Gráfico 4.3: Número de Sinistros.....	31
Gráfico 4.4: Custo Médio .....	32
Gráfico 4.5: Frequência de Sinistralidade .....	32
Gráfico 4.6: Rácio de Sinistralidade.....	33
Gráfico 6.1: Evolução da frequência (ADAS) .....	47
Gráfico 6.2: Evolução do custo médio (ADAS).....	47

## Lista de Figuras

Figura 2.1: Tabela Bonús-Malus .....	12
Figura 5.1: Tabela inicial preenchida com a variável distrito .....	43
Figura 5.2: Tabela ANOVA e Conclusão.....	44
Figura 5.3: Conclusão sobre o número de grupos encontrado .....	44
Figura 5.4: Conclusão quando o número de grupos não é adequado .....	44
Figura 5.5: Output para a construção de grupos.....	45



# 1. Atividade Seguradora

Neste capítulo é feita uma primeira abordagem à atividade seguradora. Assim, é apresentada a história dos seguros, os principais conceitos da atividade seguradora e os ramos da atividade segura.

## 1.1. Breve história dos seguros

Recuando ao século XVII A.C., na antiga civilização da Mesopotâmia, havia um código, o código de Hammurabi, que continha disposições relativas a um sistema de assistência mútua, para o caso de as mercadorias transportadas em caravanas não chegarem ao seu destino. Na Grécia antiga e em Roma existiam também sistemas de assistência mútua. Tratava-se de algo que poderíamos designar um ‘pré-seguro’, antecedente daquilo a que hoje chamamos ‘seguro’.

O atual conceito de ‘seguro’ consiste na transferência do risco para um terceiro, que nada tem a ver com a atividade em risco, mediante o pagamento de uma quantia fixada previamente, conhecida como prémio. Surgiu na idade média com o contrato de empréstimo, conhecido como *foenus nauticum* ou como *empréstimo para a grande aventura*, que acabou por originar o seguro marítimo. Nesta época foi criado um esquema em que os comerciantes se associavam e aceitavam garantir, mediante o pagamento de um valor previamente fixado, que caso houvesse um naufrágio com prejuízo, o seu barco e a respetiva carga seriam reembolsados. O contrato escrito, conhecido como *apólice*, era rasgado no fim da viagem. O primeiro contrato nestes moldes conhecido, terá sido celebrado em 1347, em Génova.

Devido à importância económica, comércio marítimo e sistema bancário, de alguns estados italianos nos séculos XIV a XVII, Génova, Veneza e Florença foram os principais centros de atividade seguradora na Europa. O primeiro intermediário de seguros surgiu em Florença no século XIV, e a primeira empresa de seguros marítimos apareceu em Génova em 1424.

No século XVII, como consequência da decadência económica das repúblicas italianas, os principais centros de comércio internacional foram deslocados para Londres e Amesterdão, aí relocando os centros seguradores da Europa.

Foi em 1684 que foi criado o seguro contra o risco de incêndio, devido a um gigantesco e catastrófico incêndio, conhecido como o “Grande Incêndio”, ocorrido em 1666, na cidade de Londres, que destruiu muitas casas e vitimou muitas pessoas.

Em Portugal, foi durante o reinado de D. Fernando (segunda metade do século XIV), que surgiu uma espécie de mercado segurador: a *Bolsa de Seguros*. Durante o período das descobertas, nomeadamente durante o reinado de D. Manuel I (início do século XVI), começaram a surgir indivíduos que exerciam a intermediação de seguros, ligando os que se queriam proteger do risco, e os que, a troco de remuneração, aceitavam esse mesmo risco. Tal como noutros países, em Portugal, os contratos eram feitos verbalmente. De forma a existir um maior controlo, nomeadamente, de fiscalização deste negócio, foi criado, em 1529, o cargo de *Escrivão de Seguros*, o que equivale hoje à autoridade portuguesa de regulação e supervisão de seguros, conhecida por Autoridade de Supervisão de Seguros e Fundos de Pensões (ASF). Em meados do século XVI foi criada a *Casa dos Seguros*, que funcionava como central de registo e negociação de contratos.

A primeira companhia privada de seguros em Portugal acabou por não sobreviver às invasões francesas, pelo que a atividade seguradora em Portugal continuou a ser exercida por companhias seguradoras estrangeiras, maioritariamente inglesas até 1848, ano em que nasceu a companhia de seguros Fidelidade.

## 1.2. Principais Conceitos da Atividade Seguradora

Em primeiro lugar é importante perceber o que é o **seguro**. Uma definição interessante é a de Joseph Hemard, que nos diz que: *“O seguro é uma operação pela qual alguém, o segurado, obtém a promessa de que, mediante renumeração - o prêmio -, em caso de concretização do risco, ele (ou um terceiro beneficiário) será ressarcido por outrem – o segurador- que, assumindo o conjunto de riscos, os compensará de acordo com as leis estatísticas”*.

A definição de seguro, também conhecido como contrato de seguro ou apólice, inclui conceitos importantes, entre eles, o conceito de risco. Nos seguros, o **risco** é um acontecimento futuro, independente da vontade do segurado e completamente aleatório.

Pelo que vimos nas definições anteriores, numa apólice existem vários intervenientes: o segurado, o tomador de seguro e o segurador.

O **segurado** é a pessoa cujos interesses patrimoniais, a vida ou a saúde, são protegidos pelo seguro. O **tomador de seguro** é a pessoa que subscreve o seguro e que é responsável pelo pagamento do prêmio, pela declaração do risco a segurar e pela participação de eventuais sinistros. O **segurador**, é a companhia de seguros (daqui em diante denominada Companhia), que assume a responsabilidade do risco em troca de um prêmio.

Surgem assim dois conceitos importantes: o de prêmio e o de sinistro. O **prêmio** é o valor que o tomador do seguro paga à seguradora, para que esta efetue a gestão dos riscos e das coberturas que fazem parte da apólice. O **sinistro** é a efetivação de um evento previsto na apólice e que desencadeia o cumprimento das obrigações da seguradora. Ou seja, é quando o risco previsto no contrato se concretiza e origina danos.

É também importante perceber o que é uma franquía e a que corresponde o capital seguro. A **franquia** é a parte do custo de um sinistro que fica a cargo do tomador de seguro. O **capital seguro** é o valor do bem seguro quando é feito o seguro, é também o valor que serve de base para o cálculo do prêmio. É sobre esse valor que recai a responsabilidade transferida do segurado para a seguradora em caso de sinistro.

## 1.3. Ramos da atividade seguradora

A atividade seguradora divide-se em dois grandes ramos, o ramo vida e o ramo não vida.

### 1.3.1. Ramo Vida

Os seguros do ramo vida centram-se na proteção do valor da vida do segurado ou na garantia de certa rentabilidade ao segurado. Assim, este ramo divide-se nos sub-ramos que descrevemos em seguida.

#### 1.3.1.1. Seguros de Vida

Trata-se de um seguro em que o risco principal é a morte, a sobrevivência ou ambos, de uma pessoa ou várias. Garante assim o pagamento de um capital ou de uma renda ao segurado, em caso de vida, ou aos beneficiários, em caso de morte.

#### **1.3.1.2. Seguros de Nupcialidade ou Natalidade**

Trata-se de um seguro que garante o pagamento de um capital ou renda em caso de casamento ou em caso de nascimento de filhos, respetivamente.

#### **1.3.1.3. Operações de Capitalização**

São contratos em que a seguradora garante um capital fixo, definido previamente, decorrido um certo período, em troca do pagamento de um prémio único ou em prestações.

#### **1.3.1.4. Seguros Ligados a Fundos de Investimento**

Trata-se de um seguro de capital variável, em que o capital que o segurado recebe depende, na totalidade ou parcialmente, de um “valor de referência” constituído por uma ou mais “unidades de participação”. Trata-se, pois, de um seguro em que o segurado assume o risco de capital.

### **1.3.2. Ramo Não Vida**

Os seguros do ramo não vida são seguros cujo objetivo é proteger um determinado bem. Este ramo subdivide-se em:

#### **1.3.2.1. Ramo Automóvel**

Um dos principais tipos de seguro deste ramo é o seguro obrigatório de responsabilidade civil automóvel. O objetivo deste seguro é proteger qualquer veículo terrestre a motor que circule na via rodoviária contra os riscos de circulação. Existem ainda outros seguros neste ramo, de natureza opcional como, por exemplo, o seguro contra danos próprios.

#### **1.3.2.2. Ramo Transportes**

Este tipo de seguro protege qualquer veículo aéreo, marítimo ou terrestre. Neste ramo, também o seguro de responsabilidade civil é obrigatório. Pode também este seguro cobrir a mercadoria transportada por estes veículos.

#### **1.3.2.3. Ramo Responsabilidade Civil**

Neste ramo é segurado o risco de qualquer dano contra um terceiro, assegurando assim a indemnização do mesmo pela seguradora.

#### **1.3.2.4. Ramo Patrimoniais**

O ramo patrimoniais subdivide-se nos sub-ramos que descrevemos em seguida.

#### **1.3.2.4.1. Seguro Multirriscos Habitação**

O seguro multirriscos habitação cobre danos no imóvel ou no seu recheio, podendo também incluir outras coberturas, como a de responsabilidade civil. A cobertura de incêndio é obrigatória.

#### **1.3.2.4.2. Seguro Multirriscos Indústria e Comércio**

O seguro multirriscos indústria protege qualquer espaço do sector indústria, já o seguro multirriscos comércio cobre a atividade resultante do comércio, podendo cobrir o edifício e/ou o seu recheio.

#### **1.3.2.5. Ramo Acidentes**

O ramo acidentes divide-se em dois tipos de seguro distintos, nomeadamente, o de acidentes de trabalho e o de acidentes pessoais.

##### **1.3.2.5.1. Acidentes de Trabalho**

Este seguro destina-se a proteger a pessoa contra qualquer acidente que ocorra no seu horário e local de trabalho ou nos trajetos trabalho-casa e casa-trabalho. É um seguro obrigatório para trabalhadores por conta de outrem e para trabalhadores independentes.

##### **1.3.2.5.2. Acidentes Pessoais**

Este seguro destina-se a proteger a pessoa contra qualquer acidente que ocorra no seu dia-a-dia e que não esteja abrangido pelo seguro automóvel ou pelo seguro de acidentes de trabalho. Pode ser realizado para um período específico, como para uma viagem.

#### **1.3.2.6. Ramo Saúde**

O seguro de saúde serve para proteger a pessoa contra o risco de determinadas doenças e para cobrir cuidados de saúde, como consultas de especialidade, medicamentos, óculos e outros tipos de assistência médica.



## **2. Seguro Automóvel**

Uma vez que o objetivo deste trabalho se insere no ramo automóvel, este capítulo apresenta mais detalhadamente alguns conceitos relacionados com este ramo, nomeadamente, os diferentes tipos de seguro, o fundo de garantia automóvel, o sistema de indemnização direta ao segurado e a tabela bônus-malus.

### **2.1. Tipos de seguro**

No seguro automóvel existem dois seguros principais - responsabilidade civil e danos próprios - e um seguro que cobre outros danos, que pode ser subscrito em conjunto com qualquer um dos principais.

#### **2.1.1. Responsabilidade Civil**

É um seguro, obrigatório, que garante o pagamento das indemnizações por danos materiais e corporais causados a terceiros e aos ocupantes do veículo, exceto ao condutor do mesmo. Neste seguro, atualmente, o capital mínimo obrigatório é de 7.290.000€, podendo ser, opcionalmente, estendido para os 50.000.000€.

#### **2.1.2. Danos Próprios**

Este seguro visa proteger o condutor contra danos materiais que ocorram num sinistro, desde que este seja o responsável pelo mesmo. Caso contrário, o sinistro será coberto pelo seguro de responsabilidade civil do outro veículo. Dentro da cobertura de danos próprios estão as coberturas:

##### **➤ Choque, Capotamento e Colisão**

Cobre o risco de choque (embate contra um objeto ou veículo estático), colisão (embate contra um objeto ou veículo em movimento) e capotamento (quando um veículo perde a sua posição habitual e fica impossibilitado de circular).

##### **➤ Incêndio, Raio ou Explosão**

Cobre o risco de danos no veículo decorrentes de um incêndio, da queda de um raio ou de uma explosão.

##### **➤ Furto ou Roubo**

Cobre o risco de furto ou roubo do veículo ou de objetos no seu interior. A diferença entre furto e roubo, é que roubo implica violência por parte do assaltante e furto não implica.

#### **2.1.3. Outros Danos**

Existem outros riscos que não se encontram cobertos pelos seguros descritos anteriormente. O seguro de outros danos é um seguro com coberturas facultativas que pode ser subscrito para cobrir:

➤ Proteção Jurídica

Cobre as despesas com o advogado que representa o segurado e as despesas decorrentes do processo judicial.

➤ Assistência em Viagem

Cobre, em caso de avaria ou acidente, o reboque do veículo e o transporte das pessoas e respetivos bens até à conclusão da viagem.

➤ Atos de Vandalismo

Cobre os danos materiais no veículo, ocorridos como consequência de atos de vandalismo ou atos maliciosos ocorridos na via pública.

➤ Fenómenos Naturais

Cobre os danos materiais no veículo, ocorridos como consequência de fenómenos naturais e incêndios resultantes dos mesmos.

➤ Quebra Isolada de Vidros

Cobre o risco de quebra ou rotura de qualquer vidro existente no veículo, garantindo assim a indemnização das despesas de substituição e montagem dos mesmos.

➤ Privação de Uso

Garante uma indemnização ao segurado das despesas decorrentes da impossibilidade de utilizar o seu veículo decorrente de um sinistro.

➤ Acidentes Pessoais Ocupantes

Garante o pagamento de capitais ou indemnizações por lesões sofridas pelos ocupantes do veículo.

## **2.2. Fundo de garantia automóvel**

O fundo de garantia automóvel é um fundo que garante, em caso de acidente, a indemnização das vítimas, caso o responsável pelo acidente não tenha um seguro válido ou caso este seja desconhecido. Este fundo cobre indemnizações por morte ou por danos corporais. Apenas cobre danos materiais quando o responsável é conhecido, mas não apresenta um seguro válido. Todas as apólices pagam uma contribuição para este fundo.

## **2.3. Sistema de regularização de sinistros IDS**

A Indemnização Direta ao Segurado (IDS) é um sistema de regularização de sinistros que permite que seja a seguradora do lesado a reparar o veículo ao seu cliente, isto é, permite que, em caso de sinistro, mesmo que o segurado não seja culpado, seja a sua seguradora a pagar-lhe os custos com o sinistro. Este sistema funciona como uma parceria entre seguradoras, em que cada seguradora paga ao seu próprio segurado. A seguradora recebe depois o reembolso do valor correspondente ao valor médio dos processos IDS. Desta forma os sinistros de responsabilidade civil são processados de forma mais rápida. Só se pode aplicar este sistema se:

- Envolver apenas dois veículos;
- Tiver havido uma colisão entre eles;
- Tiverem seguros de seguradoras aderentes ao protocolo IDS;
- O sinistro tiver ocorrido em Portugal;
- Os danos materiais em cada veículo não ultrapassarem os 15.000€;
- Existir uma declaração amigável preenchida corretamente;
- Não existirem danos corporais.

## **2.4. Bónus-Malus**

O sistema de Bónus-Malus é a forma que as seguradoras têm para compensar ou penalizar o segurado pelo comportamento que apresentou na anuidade anterior. Assim Bónus corresponde a uma diminuição percentual no valor do prémio, em algumas coberturas, se o segurado não tiver sinistros em que seja responsável. Malus corresponde a um aumento percentual no valor do prémio, de algumas coberturas, se o segurado tiver sinistros da sua responsabilidade. Os clientes podem contratar uma cobertura de proteção de Bónus e assim não terão um agravamento de prémio derivado do primeiro sinistro da anuidade.

Este sistema só é aplicado para sinistros que afetem alguma das seguintes coberturas:

- Responsabilidade Civil;
- Furto ou Roubo;
- Incêndio, Raio ou Explosão;
- Danos Acidentais Sofridos pelo Veículo.

Na figura 2.1 é apresentada a tabela Bónus-Malus em vigor na Companhia. Todos os clientes que iniciam no mercado segurador iniciam na letra O, caso contrário mantêm o nível de experiência (letra bónus-malus correspondente) que trazem da companhia anterior, até ao máximo de 5 anos. Caso não tenham sinistros, na anuidade seguinte mudam para o nível N. Caso os clientes tenham sinistros pioram o nível. Esta tabela aplica-se apenas a novos clientes, pois quando há alterações na tabela os clientes ficam com os valores da tabela apresentada no seu 1º contrato.

Escalaão	Nº de Anos	% Desconto ou Agravamento	Nº de Sinistros na Última Anuidade									
			Sem Sinistro		1 Sinistro		2 Sinistros		3 Sinistros		4 Sinistros	
A	14	-57	-57	-A	-53	-E	-45	-I	-40	-K	0	-O
B	13	-55	-57	-A	-49	-G	-40	-K	-35	-L	0	-O
C	12	-55	-55	-B	-47	-H	-40	-K	-25	-M	0	-O
D	11	-55	-55	-C	-45	-I	-40	-K	-25	-M	0	-O
E	10	-53	-55	-D	-45	-I	-40	-K	-25	-M	0	-O
F	9	-51	-53	-E	-40	-K	-35	-L	-25	-M	+15	-P
G	8	-49	-51	-F	-40	-K	-25	-M	-15	-N	+15	-P
H	7	-47	-49	-G	-35	-L	-25	-M	-15	-N	+15	-P
I	6	-45	-47	-H	-35	-L	-25	-M	0	-O	+15	-P
J	5	-45	-45	-I	-35	-L	-25	-M	0	-O	+25	-Q
K	4	-40	-45	-J	-25	-M	0	-O	0	-O	+35	-R
L	3	-35	-40	-K	-15	-N	0	-O	+25	-Q	+35	-R
M	2	-25	-35	-L	0	-O	0	-O	+35	-R	+45	-S
N	1	-15	-25	-M	+15	-P	+25	-Q	+45	-S	+80	-T
O	0	0	-15	N	+25	-Q	+35	-R	+80	-T	+80	-T
P		+15	0	-O	+35	-R	+45	-S	+80	-T	+80	-T
Q		+25	+15	-P	+45	-S	+80	-T	+80	-T	+80	-T
R		+35	+25	-Q	+80	-T	+80	-T	+80	-T	+80	-T
S		+45	+35	-R	+80	-T	+80	-T	+80	-T	+80	-T
T		+80	+45	-S	+80	-T	+80	-T	+80	-T	+80	-T

Figura 2.1: Tabela Bonús-Malus

### 3. Métodos Estatísticos Utilizados

Neste capítulo são apresentados os métodos estatísticos que foram utilizados nas análises efetuadas.

#### 3.1. Intervalos de confiança

##### 3.1.1. Intervalo de confiança para o valor médio, variância desconhecida

Suponha-se que temos uma amostra  $x = (x_1, x_2, \dots, x_n)$  extraída de uma população  $X$  com distribuição *Normal*  $(\mu, \sigma)$ , em que quer o valor médio  $(\mu)$ , quer a variância  $(\sigma^2)$  são desconhecidos. O objetivo é estimar o valor médio e usamos o desvio padrão empírico  $(s)$  como estimador da variabilidade intrínseca  $(\sigma)$ .

Ao dividir  $\bar{X}_n - \mu$  pelo seu erro padrão,  $\frac{s}{\sqrt{n}}$ , estamos a *studentizar* a média empírica, isto é, estamos a dividir pelo valor observado da variável aleatória,  $\frac{s}{\sqrt{n}}$ , que é proporcional à raiz de uma variável com distribuição qui-quadrado com  $n-1$  graus de liberdade,  $\chi_{n-1}^2$ , dividida pelo seu número de graus de liberdade e, por isso, obtém-se o valor observado de uma variável com distribuição t-de-student com  $n-1$  graus de liberdade,  $t_{n-1}$ .

A variável aleatória  $\frac{\bar{X}_n - \mu}{s/\sqrt{n}}$  é simétrica e, sendo uma  $t_{n-1}$ , a sua distribuição não depende funcionalmente de  $\mu$ , ou seja, é uma variável fulcral.

De onde sai que,

$$P\left(t_{n-1, \frac{\alpha}{2}} < \frac{\mu - \bar{X}_n}{\frac{s}{\sqrt{n}}} < t_{n-1, 1-\frac{\alpha}{2}}\right) = 1 - \alpha \quad (3.1)$$

Ora como  $t_{n-1, \frac{\alpha}{2}} = -t_{n-1, 1-\frac{\alpha}{2}}$ , então também podemos escrever:

$$P\left(\bar{X}_n - t_{n-1, \frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{X}_n + t_{n-1, 1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right) = 1 - \alpha \quad (3.2)$$

Pelo que o intervalo de confiança para  $\mu$ , com nível de confiança  $1-\alpha$  é dado por:

$$\left(\bar{x}_n - t_{n-1, \frac{\alpha}{2}} \frac{s}{\sqrt{n}}, \bar{x}_n + t_{n-1, 1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right) \quad (3.3)$$

##### 3.1.2. Intervalo de confiança para uma proporção

O parâmetro  $p$  do modelo binomial é uma proporção, cujo estimador de máxima verosimilhança é dado por  $\hat{p} = \frac{\bar{X}}{n}$ .

Consequentemente, por via do teorema do limite central,  $\hat{p}$  segue assintoticamente uma distribuição normal, com valor médio  $p$  e desvio padrão  $\sqrt{\frac{p(1-p)}{n}}$ , isto é,  $\hat{p} \sim N(p, \sqrt{\frac{p(1-p)}{n}})$ .

Obtemos assim o seguinte intervalo de confiança para  $p$ , que assintoticamente tem probabilidade igual  $1-\alpha$ ,

$$(\hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}, \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}) \quad (3.4)$$

No entanto, este não é ainda um intervalo de confiança para o parâmetro, uma vez que os extremos do intervalo dependem de  $p$ , que é o parâmetro que queremos estimar. Mas, uma vez que o estimador de máxima verosimilhança de  $p$ ,  $\hat{p}$ , é consistente para o parâmetro, podemos usar o intervalo de confiança de *Wald* que se obtém substituindo o parâmetro por esse estimador. Ou seja, o intervalo de confiança assintótico para  $p$ , com um nível de confiança  $1-\alpha$  é dado por:

$$(\hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}) \quad (3.5)$$

## 3.2. Modelos de Regressão Linear

### 3.2.1. Regressão Linear Simples

Um modelo de regressão permite-nos exprimir uma variável dependente,  $y$ , em função de uma variável independente,  $x$ , ou seja,  $y=f(x)$ .

Por outras palavras, consideramos que os valores observados,  $y_i$ , são variações amostrais em torno de um modelo  $\hat{y}_i = f(x_i)$ , ou seja, os valores observados podem ser escritos como  $y_i = \hat{y}_i + e_i$ , em que  $e_i$  são os resíduos.

O modelo de regressão linear simples é dado por:

$$y_i = ax_i + b + e_i \quad (3.6)$$

Sendo que  $y$  é a variável dependente,  $x$  é a variável independente,  $a$  e  $b$  são os coeficientes desconhecidos do modelo,  $i \in \{1, \dots, n\}$ , em que  $n$  é a dimensão da amostra e  $e_i$  são os resíduos.

O método mais usual para estimar os parâmetros é o método dos mínimos quadrados, em que o objetivo é minimizar a soma dos quadrados dos desvios entre os valores observados ( $y_i$ ) e os valores estimados ( $\hat{y}_i = f(x_i)$ ). Por outras palavras, o objetivo é minimizar a soma dos quadrados dos resíduos, uma vez que  $\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ .

Queremos então determinar os coeficientes  $a$  e  $b$  tais que o desvio quadrado global  $Q(a,b) = \sum_{i=1}^n (y_i - a * x_i - b)^2$  seja mínimo.

Os estimadores de mínimos quadrados obtêm-se igualando a zero as derivadas parciais de  $Q(a,b)$  em ordem às incógnitas  $a$  e  $b$ , isto é:

$$\begin{cases} \frac{\partial}{\partial a} Q(a,b) = 0 \\ \frac{\partial}{\partial b} Q(a,b) = 0 \end{cases} \quad (3.7)$$

o que é equivalente a,

$$\begin{cases} \frac{\partial}{\partial a} Q(a, b) = -2 \sum_{i=1}^n x_i (y_i - ax_i - b) = 0 \\ \frac{\partial}{\partial b} Q(a, b) = -2 \sum_{i=1}^n (y_i - ax_i - b) = 0 \end{cases} \quad (3.8)$$

Obtemos assim um sistema de duas equações a duas incógnitas, a e b,

$$\begin{cases} \sum_{i=1}^n x_i (y_i - ax_i - b) = 0 \\ \sum_{i=1}^n (y_i - ax_i - b) = 0 \end{cases} \quad (3.9)$$

Equivalentemente temos,

$$\begin{cases} \sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i = 0 \\ \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - nb = 0 \end{cases} \quad (3.10)$$

Este sistema é ainda equivalente a este outro:

$$\begin{cases} \sum_{i=1}^n x_i y_i = a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i \quad (1) \\ \sum_{i=1}^n y_i = a \sum_{i=1}^n x_i + nb \quad (2) \end{cases} \quad (3.11)$$

Reparando que a equação (2) pode ser escrita como

$$\sum_{i=1}^n y_i - a \sum_{i=1}^n x_i = nb \quad (3.12)$$

vem que  $\hat{b}$  é igual a:

$$\hat{b} = \frac{\sum_{i=1}^n y_i}{n} - \hat{a} \frac{\sum_{i=1}^n x_i}{n} \quad (3.13)$$

Sabemos no entanto que  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$  e  $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ , logo  $\hat{b} = \bar{y} - \hat{a}\bar{x}$ .

Substituindo em (1) b por  $\hat{b} = \frac{\sum_{i=1}^n y_i}{n} - \hat{a} \frac{\sum_{i=1}^n x_i}{n}$ , obtemos:

$$\sum_{i=1}^n x_i y_i = \hat{a} \sum_{i=1}^n x_i^2 + \left( \frac{\sum_{i=1}^n y_i}{n} - \hat{a} \frac{\sum_{i=1}^n x_i}{n} \right) * \sum_{i=1}^n x_i \quad (3.14)$$

o que é equivalente a termos:

$$\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i * \sum_{i=1}^n y_i}{n} = \hat{a} \sum_{i=1}^n x_i^2 - \hat{a} \frac{\sum_{i=1}^n x_i^2}{n} \quad (3.15)$$

Se multiplicarmos e dividirmos por n, as parcelas da equação:  $\sum_{i=1}^n x_i y_i$  e  $\sum_{i=1}^n x_i^2$ , sabemos que a equação acima é equivalente a

$$\frac{n * \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n} = \hat{a} \left( \frac{n(\sum_{i=1}^n x_i^2) - \sum_{i=1}^n x_i^2}{n} \right) \quad (3.16)$$

De onde concluímos que

$$\hat{a} = \frac{n * \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n(\sum_{i=1}^n x_i^2) - \sum_{i=1}^n x_i^2} \quad (3.17)$$

Se substituirmos agora a expressão na equação (2) pela equação 3.17, chegamos a conclusão de que

$$\hat{b} = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2 (\sum_{i=1}^n x_i)^2} \quad (3.18)$$

Ou seja, pelo método dos mínimos quadrados, os estimadores para os coeficientes a e b são obtidos por:

$$\begin{cases} \hat{a} = \frac{n * \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n(\sum_{i=1}^n x_i^2) - \sum_{i=1}^n x_i^2} \\ \hat{b} = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2 (\sum_{i=1}^n x_i)^2} \end{cases} \quad (3.19)$$

$$\Leftrightarrow \begin{cases} \hat{a} = \frac{n * \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n(\sum_{i=1}^n x_i^2) - \sum_{i=1}^n x_i^2} \\ \hat{b} = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2 (\sum_{i=1}^n x_i)^2} \end{cases}$$

Conseguimos assim saber como obtemos os coeficientes para ajustar a reta de mínimos quadrados através da equação  $y_i = ax_i + b + e_i$ .

Para sabermos se a reta se ajusta bem ao conjunto de observações usamos o coeficiente de determinação, usualmente designado por  $R^2$ , que corresponde a  $R^2 = \frac{\sum_{i=1}^n (\hat{y} - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ . Este coeficiente varia entre 0 e 1, sendo que quando é igual ou muito próximo de 1 significa que grande parte da variabilidade de  $y$  é explicada por  $x$ , ou seja, corresponde a um bom ajustamento do modelo de regressão linear às observações.

### 3.2.2. Regressão Linear Múltipla

Sejam  $X_1, X_2, \dots, X_k$  as variáveis explicativas,  $\beta_1, \beta_2, \dots, \beta_k$  os coeficientes do modelo e  $\varepsilon$  os erros aleatórios. O modelo de regressão múltipla (RLM) é uma generalização do modelo de regressão linear simples em que a variável resposta  $y$  se escreve como uma função das  $k$  variáveis explicativas e pode então ser escrito como:

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k + \varepsilon \quad (3.20)$$

Considere-se uma amostra de dimensão  $N$  deste modelo,

$$\{(y_i, x_{i1}, x_{i2}, \dots, x_{ik}): i=1, 2, \dots, N\}$$

A partir do modelo (3.20) obtém-se:

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_k x_{ik} + \varepsilon_i, i=1, \dots, N \quad (3.21)$$

As  $N$  igualdades (3.21) podem ser apresentadas em notação matricial, pois considerando

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix}, X = \begin{bmatrix} 1 & x_{12} & x_{13} & \dots & x_{1k} \\ 1 & x_{22} & x_{23} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{N2} & x_{N3} & \dots & x_{Nk} \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_N \end{bmatrix},$$

podemos escrever o modelo de RLM como

$$Y = X\beta + \varepsilon \quad (3.22)$$

Para estimar os coeficientes de regressão,  $\beta$ , continua a utilizar-se o método dos mínimos quadrados (MMQ). Os estimadores de MQ destes parâmetros obtêm-se minimizando a soma dos quadrados dos resíduos,

$$\sum_{i=1}^N \varepsilon_i^2 = \sum_{i=1}^N [y_i - (\beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_k x_{ik})]^2 \quad (3.23)$$

Derivando esta expressão em ordem a cada um dos coeficientes e após alguns cálculos matriciais obtemos os estimadores de MQ dos coeficientes, que são dados por



$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (3.24)$$

onde  $X^T$  representa a matriz transposta de  $X$ . Tem-se ainda que a matriz de covariâncias dos estimadores dos mínimos quadrados é dada por

$$V(\hat{\beta}) = \sigma^2 (X^T X)^{-1} \quad (3.25)$$

e um estimador centrado dos resíduos é dado por

$$\hat{\sigma}^2 = \left( \frac{1}{N-k} \right) \sum_{i=1}^n \hat{\varepsilon}_i^2 \quad (3.26)$$

Pode-se ainda demonstrar que  $\hat{\beta}$  segue assintoticamente uma distribuição normal k-dimensional com parâmetros  $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ , valores esperados dos estimadores  $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)$  e matriz de variâncias e covariâncias  $\sigma^2 (X^T X)^{-1}$ .

### 3.3. Análise de Variância Simples

#### 3.3.1. Teste de Igualdade de Médias

O teste de igualdade de médias permite-nos comparar duas médias de forma a sabermos se existem diferenças significativas entre as mesmas.

Consideremos duas amostras independentes,  $(X_{11}, \dots, X_{1n_1})$  e  $(X_{21}, \dots, X_{2n_2})$  de dimensão  $n_1$  e  $n_2$ , provenientes de populações com distribuição normal com média  $\mu_1$  e  $\mu_2$ , e com variância comum  $\sigma^2$ . Nestes pressupostos, pretendemos testar as hipóteses

$$H_0: \mu_1 = \mu_2 \text{ vs } H_1: \mu_1 \neq \mu_2 \quad (3.27)$$

Em determinadas situações, podemos fazer um teste unilateral em que a hipótese alternativa é da forma  $H_1: \mu_1 > \mu_2$  ou  $H_1: \mu_1 < \mu_2$ .

Nestas condições, a diferença de médias segue também uma distribuição normal, ou seja,

$$\bar{X}_1 - \bar{X}_2 \cap N(\mu_1 - \mu_2, \sigma^2 (\frac{1}{n_1} + \frac{1}{n_2})) \quad (3.28)$$

Um bom estimador da variância comum às duas populações é dado por

$$S_T^2 = \frac{1}{n_1 + n_2 - 2} \sum_{j=1}^2 \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2 \quad (3.29)$$

Este estimador é tal que

$$\frac{(n_1 + n_2 - 2)}{\sigma^2} S_T^2 \cap \chi_{n_1 + n_2 - 2}^2 \quad (3.30)$$

e é independente de  $\bar{X}_1 - \bar{X}_2$ . Assim sendo, a estatística de teste é dada por

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_T \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (3.31)$$

que, nas condições já referidas e sob a validade de  $H_0$ , segue uma distribuição t-student com  $n_1 + n_2 - 2$  graus de liberdade.

### 3.3.2. Análise de Variância

Vimos no subcapítulo anterior que o teste de igualdade de médias pode ser utilizado para comparar duas médias, mas se quisermos comparar mais do que duas médias já não o podemos utilizar. Nessa situação podemos utilizar um modelo de análise de variância, também conhecida por ANOVA.

Neste caso pretendemos testar as hipóteses

$$H_0: \mu_1 = \mu_2 = \dots = \mu_n \text{ vs } H_1: \exists \text{ pelo menos um } \mu_k \text{ diferente}, k = 1, \dots, n \quad (3.32)$$

Vamos começar por ver como exprimir a variabilidade das médias das  $n$  populações em torno da média global, ou seja, da média comum às  $n$  populações. Esta variabilidade corresponde à variabilidade entre amostras e é dada por

$$SQ_{ext} = \sum_{i=1}^n n_i (\bar{X}_i - \bar{X})^2 \quad (3.33)$$

em que

$$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} \text{ e } \bar{X} = \frac{1}{N} \sum_{i=1}^n \sum_{j=1}^{n_i} X_{ij}, N = \sum_{i=1}^n n_i$$

Por outro lado, a variabilidade dentro das amostras é dada por

$$SQ_e = \sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 = (N - 1)S_T^2 \quad (3.34)$$

Quando somamos a variabilidade entre amostras com a variabilidade dentro da amostra obtemos a variabilidade total da amostra, ou seja,

$$SQ_{tot} = SQ_e + SQ_{ext} \quad (3.35)$$

em que a variabilidade total é dada por:

$$SQ_{tot} = \sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2 \quad (3.36)$$

Para comprovarmos que  $SQ_{tot} = SQ_e + SQ_{ext}$ , vamos somar e subtrair a média de cada um dos grupos na fórmula de  $SQ_{tot}$ , ou seja,

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2 &= \sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} + \bar{X}_i - \bar{X}_i - \bar{X})^2 = \sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 + \\ &+ \sum_{i=1}^n n_i (\bar{X}_i - \bar{X})^2 + 2 * \sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}) \end{aligned} \quad (3.37)$$

Pode-se ver facilmente que o termo retangular é nulo, já que se tem

$$\sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}) = \sum_{i=1}^n \left[ (\bar{X}_i - \bar{X}) \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i) \right] = \sum_{i=1}^n (\bar{X}_i - \bar{X}) (n_i \bar{X}_i - n_i \bar{X}_i) = 0 \quad (3.38)$$

Logo, verifica-se a igualdade

$$\sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2 = \sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 + \sum_{i=1}^n n_i (\bar{X}_i - \bar{X})^2 \quad (3.39)$$

Se considerarmos que o estimador da variância da amostra  $i$  é dado por

$$S_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 \quad (3.40)$$

vem que as variáveis  $\frac{(n_i-1)}{\sigma^2} S_i^2$  são independentes, uma vez que são obtidas a partir de amostras independentes, cada uma com distribuição qui-quadrado com  $n_i-1$  graus de liberdade. Pelo que a sua soma tem distribuição qui-quadrado com  $N-n$  graus de liberdade.

Se a hipótese nula,  $H_0$ , for verdadeira, podemos considerar o conjunto das  $n$  amostras como uma amostra única proveniente de uma população com distribuição normal. Assim  $\frac{SQ_{tot}}{N-1}$  é um estimador centrado para essa variância, e a variável  $\frac{SQ_{tot}}{\sigma^2}$  tem distribuição qui-quadrado com  $N-1$  graus de liberdade.

$SQ_e$  e  $SQ_{ext}$  são independentes, uma vez que  $SQ_e$  depende da variância de cada amostra e  $SQ_{ext}$  depende apenas das médias das mesmas amostras. Para amostras recolhidas de populações normais, a variância da amostra é independente da respetiva média, pelo que podemos concluir que as duas quantidades são independentes. Podemos então escrever a equação da variabilidade como

$$\frac{SQ_{tot}}{\sigma^2} = \frac{SQ_e}{\sigma^2} + \frac{SQ_{ext}}{\sigma^2} \quad (3.41)$$

o que, em termos de distribuições de probabilidade, pode ser escrito como

$$\chi_{N-1}^2 = \chi_{N-n}^2 + \chi_{n-1}^2 \quad (3.42)$$

Assim, podemos definir a estatística de teste como sendo

$$F = \frac{\frac{SQ_{ext}}{n-1}}{\frac{SQ_e}{N-n}} \quad (3.43)$$

Esta estatística, sob a validade de  $H_0$ , corresponde ao quociente de duas variáveis independentes com distribuição qui-quadrado divididas pelo respetivo número de graus de liberdade. Deste modo, a estatística tem distribuição  $F$  com  $n-1$  graus de liberdade no numerador e  $N-n$  no denominador. Esta estatística pode também ser apresentada como

$$F = \frac{MQ_{ext}}{MQ_e} \quad (3.44)$$

em que  $MQ$  se refere a média de quadrados que, neste caso, são dadas por

$$MQ_{ext} = \frac{SQ_{ext}}{n-1} \text{ e } MQ_e = \frac{SQ_e}{N-n} \quad (3.45)$$

O resultado do teste é normalmente apresentado na forma de tabela, conhecida como a tabela da ANOVA. A estrutura da tabela é apresentada em baixo.

Fonte de Variação	Soma de quadrados	Graus de liberdade	Média de quadrados	Estatística F
Entre grupos	$\sum_{i=1}^n n_i * (\bar{X}_i - \bar{X})^2$	$n-1$	$MQ_{ext} = \frac{SQ_{ext}}{n-1}$	$F = MQ_{ext}/MQ_e$
Erro	$\sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$	$N-n$	$MQ_e = \frac{SQ_e}{N-n}$	
Total	$\sum_{i=1}^n \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$	$N-1$	$MQ_e = \frac{SQ_{tot}}{N-1}$	

### 3.3.3. Análise de variância como modelo de regressão

No subcapítulo anterior vimos como generalizar o teste de igualdade de médias a  $n$  populações normais com base em  $n$  amostras independentes entre si,  $(X_{i1}, X_{i2}, \dots, X_{in_i}), i=1, \dots, n$ , em que a amostra  $i$  é proveniente de uma população  $X_i$  com distribuição normal. Podemos assim dizer que cada observação pode ser escrita como

$$X_{ij} = \mu_i + \varepsilon_{ij}, i = 1, \dots, n, j = 1, \dots, n_i \quad (3.46)$$

em que os  $\varepsilon_{ij}$  são variáveis aleatórias independentes e com distribuição normal, com valor médio 0 e variância  $\sigma^2$ .

Vimos também que as hipóteses que queremos testar são

$$H_0: \mu_1 = \mu_2 = \dots = \mu_n \text{ vs } H_1: \exists \text{ pelo menos um } \mu_k \text{ diferente}, k = 1, \dots, n \quad (3.47)$$

Iremos agora ver que o modelo de análise de variância pode também ser escrito como um modelo de regressão linear múltipla.

Para tal, consideremos o vetor de observações,  $Y$ , na forma

$$Y' = [Y_{11} \dots Y_{1n_1} | Y_{21} \dots Y_{2n_2} | Y_{n1} \dots Y_{nn_n}] \quad (3.48)$$

em que  $\sum_{i=1}^n n_i = N$ . A matriz de planeamento pode ser escrita recorrendo a variáveis binárias,  $X_{ij}$ , isto é, variáveis que tomam apenas o valor 1 ou 0, conforme a observação pertença ou não à população  $j$ ,  $j=1, \dots, n$ , ou seja,

$$X_{ij} = \begin{cases} 1, & i = n_1 + \dots + n_{j-1} + 1 + \dots + n_j \\ 0, & \text{caso contrário} \end{cases} \quad (3.49)$$

Pelo que a matriz de planeamento é uma matriz em que os elementos assumem apenas os valores 0 ou 1, em que a  $j$ -ésima coluna está associada à  $j$ -ésima população, assumindo o valor 1 na  $i$ -ésima linha se a observação  $Y_i$  pertencer a população  $j$  e 0 caso contrário, ou seja, a matriz de planeamento toma a forma

$$X = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (3.50)$$

O vetor dos erros,  $\varepsilon_{ij}$ , é também apresentado na forma

$$\varepsilon' = [\varepsilon_{11} \dots \varepsilon_{1n_1} | \varepsilon_{21} \dots \varepsilon_{2n_2} | \dots | \varepsilon_{n1} \dots \varepsilon_{nn_n}] \quad (3.51)$$

Concluimos assim que o modelo de análise da variância pode ser escrito como um modelo de regressão, isto é,

$$Y = X\mu + \varepsilon \quad (3.52)$$

em que o vetor dos coeficientes de regressão,  $\mu$ , é constituído pelas médias das  $n$  populações,  $\mu' = [\mu_1 \mu_2 \dots \mu_n]$ . Neste caso a matriz  $(X'X)$  é dada por

$$X'X = \begin{bmatrix} n_1 & 0 & \dots & 0 \\ 0 & n_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & n_n \end{bmatrix} \quad (3.53)$$

e a matriz  $(X'Y)$  é dada por

$$X'Y = \begin{bmatrix} \sum_{j=1}^{n_1} Y_{1j} \\ \sum_{j=1}^{n_2} Y_{2j} \\ \vdots \\ \sum_{j=1}^{n_n} Y_{nj} \end{bmatrix} \quad (3.54)$$

Assim os estimadores de mínimos quadrados para os coeficientes de regressão, são os valores médios das  $n$  populações e são dados por

$$\hat{\mu} = \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \vdots \\ \hat{\mu}_n \end{bmatrix} = (X'X)^{-1}(X'Y) = \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \\ \vdots \\ \bar{Y}_n \end{bmatrix} \quad (3.55)$$

em que  $\bar{Y}_i$  corresponde à média amostral da população  $i$ . Os estimadores de mínimos quadrados dos valores médios de cada população podem também ser obtidos por derivação da soma de quadrados, que é dada por

$$SQ = \sum_{i=1}^n \sum_{j=1}^{n_i} (Y_{ij} - \mu_i)^2 \quad (3.56)$$

Assim, para cada  $k=1, \dots, n$ , temos

$$\frac{\partial SQ}{\partial \mu_k} = -2 * \sum_{j=1}^{n_k} (Y_{kj} - \mu_k) = 0 \quad (3.57)$$

o que é equivalente a

$$\sum_{j=1}^{n_k} Y_{kj} = n_k * \mu_k \quad (3.58)$$

De modo que os estimadores de mínimos quadrados são dados por

$$\widehat{\mu}_k = \bar{Y}_k = \frac{1}{n_k} \sum_{j=1}^{n_k} Y_{kj} \quad (3.59)$$

exatamente como já tínhamos deduzido a partir da fórmula dos estimadores de mínimos quadrados para um modelo de regressão em forma matricial.

Como  $Y_{ij} = \mu_i + \varepsilon_{ij}$ , imediatamente concluímos que os valores ajustados são dados por

$$\widehat{Y}_{ij} = \widehat{\mu}_i = \bar{Y}_i \quad (3.60)$$

Neste contexto, a fórmula da decomposição da soma dos quadrados diz-nos que

$$\begin{aligned} SQ_{tot} &= \sum_{i=1}^n \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2 = \sum_{i=1}^n \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 + \sum_{i=1}^n n_i * (\bar{Y}_i - \bar{Y})^2 \\ &= SQ_e + SQ_{ext} \end{aligned} \quad (3.61)$$

em que  $\bar{Y} = \frac{1}{N} * \sum_{i=1}^n \sum_{j=1}^{n_i} Y_{ij}$ ,  $N = \sum_{i=1}^n n_i$ .

Com esta formulação o teste de igualdade de médias pode ser obtido como um caso particular da hipótese linear, isto é, pode ser formulado como um teste da forma

$$H_0: C\mu = 0 \text{ contra } H_1: C\mu \neq 0 \quad (3.62)$$

em que C é uma matriz n-1 por n dada por

$$C = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix} \quad (3.63)$$

Com o objectivo de obtermos a estatística de teste, consideremos  $SQ_{tot}(C)$ ,  $SQ_e(C)$  e  $SQ_{reg}(C)$ , respetivamente, as somas de quadrados total, do erro e da regressão para o modelo completo, isto é, sem verificar a hipótese linear; e sejam  $SQ_{tot}(R)$ ,  $SQ_e(R)$  e  $SQ_{reg}(R)$ , respetivamente, as somas de quadrados total, do erro e da regressão para o modelo reduzido, ou seja, o modelo que verifica a hipótese linear. Então a estatística de teste é dada por:

$$F = \frac{[(SQ_e(R) - SQ_e(C))/q]}{[SQ_e(C)/(N-n)]} \quad (3.64)$$

e esta estatística tem distribuição F com q graus de liberdade no numerador e n-p no denominador em que q é o número de restrições, ou seja, o número de linhas da matriz C, e p é o número de variáveis.

Recorrendo à notação já estabelecida e aos resultados já apresentados sobre a hipótese linear podemos facilmente deduzir a estatística de teste para as hipóteses

$$H_0: \mu_1 = \dots = \mu_n \text{ vs } H_1: \mu_i \neq \mu_j \text{ para algum } i \neq j. \quad (3.65)$$

Sob a validade de  $H_0$ , as observações  $Y_{ij}$  têm média comum  $\mu$  e variância  $\sigma^2$ . Assim, pode-se facilmente ver que, nessas circunstâncias, o estimador de mínimos quadrados do valor médio  $\mu$  é simplesmente a média de todas as observações. Ora como,

$$\widehat{Y}_{ij} = \hat{\mu} = \bar{Y} \quad (3.66)$$

vem que a soma de quadrado dos erros para este modelo reduzido é dada por

$$SQ_e(R) = \sum_{i=1}^n \sum_{j=1}^{n_i} (Y_{ij} - \widehat{Y}_{ij})^2 = \sum_{i=1}^n \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2 = SQ_{tot} \quad (3.67)$$

Por sua vez, para o modelo completo, ou seja, para o modelo em que os valores médios são diferentes, temos

$$SQ_e(C) = SQ_{tot} - SQ_{ext}. \quad (3.68)$$

De modo que, para esta hipótese, a diferença da soma de quadrados é dada por

$$SQ_e(R) - SQ_e(C) = SQ_{tot} - (SQ_{tot} - SQ_{ext}) = SQ_{ext}. \quad (3.69)$$

Então se a hipótese nula for verdadeira, a variável

$$\frac{SQ_e(R) - SQ_e(C)}{\sigma^2} = \frac{SQ_{ext}}{\sigma^2} \quad (3.70)$$

tem distribuição qui-quadrado com n-1 graus de liberdade. Este número de graus de liberdade corresponde ao número de parâmetros da hipótese nula.

Além disso, como a variável  $SQ_e(R) - SQ_e(C)$  é independente de  $SQ_e(C)$  que, dividida por  $\sigma^2$ , tem também distribuição qui-quadrado mas com  $N-n$  graus de liberdade. Assim, vem que a estatística de teste para esta hipótese linear é dada por

$$\frac{[SQ_e(R) - SQ_e(C)]/(n-1)}{SQ_e(C)/(N-n)} = \frac{SQ_{ext}/(n-1)}{SQ_e/(N-n)} \quad (3.71)$$

que, sob validade de  $H_0$  tem distribuição F com  $n-1$  graus de liberdade no numerador e  $N-n$  no denominador. Pelo que podemos assim concluir que o teste que tínhamos obtido para a análise de variância pode também ser obtido como um teste numa hipótese linear num modelo de regressão múltipla.

No caso de a hipótese sobre a igualdade de médias ser rejeitada, escrever um modelo de análise de variância como um modelo de regressão torna-se muito útil, uma vez que permite dividir as populações em grupos com valores médios iguais entre si.

O modelo da ANOVA pode também ser apresentado através de um outro conjunto de parâmetros, a saber, considera-se a reparametrização do mesmo modelo dada por

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, i = 1, \dots, I, j = 1, \dots, n_i \quad (3.72)$$

em que os novos parâmetros são dados por

$$\mu = \frac{1}{N} \sum_{i=1}^I n_i \mu_i \text{ e } \alpha_i = \mu_i - \mu, i=1, \dots, I \quad (3.73)$$

e tais que

$$\sum_{i=1}^I n_i \alpha_i = \sum_{i=1}^I n_i \mu_i - N\mu = 0 \quad (3.74)$$

Como os novos parâmetros são funções lineares dos anteriores, vem que os EMQ são as mesmas funções lineares dos EMQ de  $\mu_1, \mu_2, \dots, \mu_I$ , isto é,

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^I n_i \bar{Y}_i = \frac{1}{N} \sum_{i=1}^I \sum_{j=1}^{n_i} Y_{ij} = \bar{Y} \quad (3.75)$$

e

$$\hat{\alpha}_i = \hat{\mu}_i - \hat{\mu} = \hat{Y}_i - \hat{Y} \quad (3.76)$$

Esta reparametrização do modelo envolve  $I+1$  parâmetros, nomeadamente,  $\mu$  e  $\alpha_1, \alpha_2, \dots, \alpha_I$ . No entanto, como  $\sum_{i=1}^I n_i \alpha_i = 0$ , existem apenas  $I$  parâmetros independentes, já que um dos  $\alpha_i$ 's pode-se escrever como combinação linear dos outros.

Com esta nova reparametrização o teste de igualdade de médias toma a forma:

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_I = 0 \text{ vs } H_1: \alpha_i \neq 0 \text{ para algum } i \quad (3.77)$$

Repare-se que, em ambos os modelos,

$$\hat{Y}_{ij} = \bar{Y}_i \quad (3.78)$$

de maneira que a partição da soma de quadrados é a mesma, ou seja, é dada por

$$SQ_{tot} = SQ_e + SQ_{ext} = SQ_e + \sum_{i=1}^I n_i \hat{\alpha}_i^2 \quad (3.79)$$

o que nos leva, naturalmente, à mesma estatística de teste F. Em ambos os casos a tabela ANOVA é a mesma e é igual a que foi apresentada no capítulo 3.3.2.

## 4. Análise da carteira automóvel da Companhia

Neste capítulo foi estudada a evolução da carteira automóvel da Companhia entre os anos 2014 e 2017. Para tal foram calculados alguns indicadores chave que nos permitem saber como a carteira se comporta quer a nível de prémios, quer a nível de sinistros. Foi também estudado o impacto das medidas que foram tomadas pela Companhia nesse período. De maneira a perceber o comportamento da Companhia no mercado foi ainda feito um estudo da quota de mercado da Companhia por distrito nos anos 2016 e 2017 e da evolução da mesma nesses dois anos.

### 4.1. Indicadores utilizados

Neste subcapítulo são apresentados os principais indicadores de performance no setor segurador, utilizados nas análises abordadas nos capítulos 4 e 5, que daqui em diante passam a ser chamados de indicadores chave.

- **Exposição:** é a proporção de tempo que uma apólice esteve em risco num determinado ano. O valor desta quantidade varia entre 0 e 1, pois é dado por:

$$\text{Exposição} = \frac{\text{nº dias em que a apólice esteve ativa na carteira}}{\text{nº dias do ano (365 ou 366)}}$$

- **Prémio Comercial:** é o prémio pago pelo cliente e que servirá para cobrir custos de regularização de sinistros, custos de distribuição, custos administrativos, custos de resseguro e remuneração do capital económico. Neste projeto, sempre que seja referido apenas prémio, corresponde a prémio comercial.
- **Prémio Anual:** é o prémio comercial, para uma anuidade, acrescido de impostos legais.
- **Prémio Adquirido:** é a parte do prémio correspondente ao período decorrido.

$$\text{Prémio adquirido} = \text{Prémio anual} * \text{Exposição}$$

- **Prémio Médio Adquirido, por apólice:**

$$\text{Prémio Médio Adquirido} = \frac{\sum \text{Prémio Adquirido}}{\sum \text{Exposição}}$$



- **Custo Médio**, por sinistro:

$$\circ \text{ Custo Médio} = \frac{\sum \text{Custos com sinistros}}{\sum n^{\circ} \text{sinistros}}$$

- **Frequência de Sinistralidade**: corresponde ao número de sinistros nos veículos em risco.

$$\circ \text{ Frequência} = \frac{\sum n^{\circ} \text{sinistros}}{\sum \text{Exposição}}$$

- **Rácio de Sinistralidade**: diz se os prémios são suficientes para fazer face aos custos com sinistros.

$$\circ \text{ Rácio de Sinistralidade} = \frac{\sum \text{custos com sinistros}}{\sum \text{Prémios Adquirido}}$$

- **Prémio de Risco**: é o prémio mínimo necessário para pagar custos com sinistros.

$$\circ \text{ Prémio de Risco} = \text{Frequência de Sinistralidade} * \text{Custo Médio}$$

## 4.2. Metodologia utilizada

A primeira parte deste projeto consistiu num estudo da evolução da carteira automóvel da Companhia entre 2014 e 2017 de forma a perceber-se onde é necessário agir para melhorar o resultado do produto. De forma a facilitar essa perceção foram calculados indicadores chave tendo por base algumas variáveis que foram consideradas relevantes, e segundo várias perspetivas.

Os indicadores chave calculados foram:

- Número de Apólices;
- Exposição;
- Prémio Médio;
- Número de Sinistros;
- Custo Total com Sinistros;
- Custo Médio;
- Desvio-padrão dos Custos;
- Frequência de Sinistralidade (ou Frequência);
- Rácio de Sinistralidade;

- Prémio de Risco.

As variáveis consideradas foram:

- Tipo de Cliente;
- Tipo de Negócio;
- Tipo de Veículo;
- Distrito;
- Marca;
- Frequência de Pagamento;
- Tipo de Combustível;
- Escala Bónus/Malus;
- Rating de Cliente;
- Rating de Agente;
- Localização Geográfica do Agente;
- Idade do Condutor vs. Idade do Segurado;
- Anos de Carta;
- Antiguidade da Apólice;
- Franquias;
- Capital Seguro.

Todos os indicadores foram calculados para as diferentes variáveis, para uma análise global da carteira, considerando uma divisão dessa análise entre com e sem danos próprios e para as coberturas de choque, capotamento e colisão, furto ou roubo e quebra isolada de vidros. Nas análises efetuadas, os dados não são os reais da Companhia, uma vez que o número de apólices, a exposição, o número de sinistros e os custos com sinistros foram multiplicados por um coeficiente.

### 4.3. Primeiras conclusões dos resultados observados

De uma primeira análise global da carteira, foram retiradas algumas conclusões, para as variáveis que descrevemos em seguida.

- Frequência de Pagamento

Tabela 4.1: Frequência de Pagamento

Frequência pagamento	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Desvio-Padrão Custos	Frequência	Rácio Sinistralidade	Prémio Risco
Anual	2014	170 909	136 335	206,8	13 804	17 774 853	1 288	1 251	10,1%	63,0%	130,4
	2015	170 523	137 006	202,3	13 572	16 871 004	1 243	1 080	9,9%	60,9%	123,1
	2016	170 838	137 560	196,5	13 521	16 431 439	1 215	1 580	9,8%	60,8%	119,4
	2017	193 244	146 602	192,5	13 846	15 976 801	1 154	1 103	9,4%	56,6%	109,0
Fracionado	2014	94 141	60 902	230,3	9 635	13 342 050	1 385	1 734	15,8%	95,1%	219,1
	2015	86 133	59 574	225,9	9 221	11 880 578	1 288	1 497	15,5%	88,3%	199,4
	2016	70 503	50 028	223,8	7 136	8 273 639	1 159	1 312	14,3%	73,9%	165,4
	2017	62 084	41 464	223,1	5 303	6 101 627	1 151	867	12,8%	66,0%	147,2

A análise da tabela 4.1 permite-nos concluir que os clientes que pagam de forma anual são um risco preferencial, uma vez que apresentam frequência e rácios de sinistralidade mais baixos, o que acaba por se traduzir num melhor resultado para a Companhia, quando comparado com os clientes que pagam de forma fracionada. Isto verifica-se devido ao facto de esta variável traduzir as condições socioeconómicas do tomador de seguro e devido a clientes que pagam de forma anual terem um maior comprometimento com a seguradora, uma vez que clientes que pagam de forma fracionada apresentam maior probabilidade de sair da Companhia a meio da vigência do contrato.

- Localização Geográfica do Agente

Tabela 4.2: Zona Geográfica do Agente

Zona geográfica	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Desvio-Padrão Custos	Frequência	Rácio Sinistralidade	Prémio Risco
Grandes Cidade, Litoral, Ilhas	2014	212 730	160 560	222,6	19 890	26 795 402	1 347	1 498	12,4%	75,0%	166,9
	2015	195 203	151 979	219,8	18 391	23 827 878	1 296	1 212	12,1%	71,3%	156,8
	2016	176 239	139 563	215,0	16 143	19 629 502	1 216	1 550	11,6%	65,4%	140,7
	2017	180 057	134 136	211,7	14 325	16 837 344	1 175	1 050	10,7%	59,3%	125,5
Sul e Interior	2014	52 320	36 677	176,8	3 549	4 321 501	1 218	1 183	9,7%	66,6%	117,8
	2015	61 454	44 601	174,1	4 402	4 923 704	1 118	1 309	9,9%	63,4%	110,4
	2016	65 102	48 025	171,2	4 514	5 075 576	1 124	1 382	9,4%	61,7%	105,7
	2017	75 271	53 930	168,3	4 823	5 241 084	1 087	1 052	8,9%	57,7%	97,2

A análise da tabela 4.2 permite concluir que os agentes do Sul e Interior apresentam melhores indicadores, nomeadamente, frequência, custo médio e rácio de sinistralidade mais baixos, o que por sua vez faz com o que o resultado seja melhor do que o correspondente aos agentes das Grandes Cidades, Litoral e Ilhas. Possivelmente deve-se a um menor número de veículos no Sul e Interior e a um eventual custo de mão-de-obra das oficinas mais baixo.

- Tipo de Cliente

Tabela 4.3: Tipo de Cliente

Tipo Cliente	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Desvio padrão do custo	Frequência	Rácio Sinistralidade	Prémio Risco
Empresa	2014	33 050	24 714	342,7	4 694	6 753 553	1 439	1 902	19,0%	79,7%	273,3
	2015	30 729	23 572	335,0	4 364	5 965 439	1 367	1 634	18,5%	75,5%	253,1
	2016	29 651	22 612	320,2	4 091	4 693 686	1 147	981	18,1%	64,8%	207,6
	2017	33 202	23 923	302,8	3 985	4 572 968	1 147	913	16,7%	63,1%	191,2
Particular	2014	232 001	172 523	195,6	18 745	24 363 351	1 300	1 362	10,9%	72,2%	141,2
	2015	225 928	173 007	192,3	18 429	22 786 143	1 236	1 171	10,7%	68,5%	131,7
	2016	211 691	164 976	187,9	16 566	20 011 392	1 208	1 566	10,0%	64,6%	121,3
	2017	222 125	164 143	184,2	15 164	17 505 459	1 154	1 069	9,2%	57,9%	106,6

A análise da tabela 4.3 permite-nos concluir que os clientes do tipo empresa apresentam piores indicadores, ou seja, frequência e rácio de sinistralidade mais elevados e, consequentemente, prémio de risco mais elevado, do que os clientes do tipo particular. Isto pode ser consequência de, na maioria das empresas, não haver um só condutor para o mesmo veículo, ou seja, existem vários perfis de risco para esse veículo. Além disso, nas empresas a frequência de utilização é tipicamente maior, pois a maioria dos particulares utiliza o carro para fazer o trajeto casa-trabalho e durante o tempo de trabalho o carro está parado, enquanto que na maioria das empresas os carros levam o dia a circular.

- Tipo de Negócio

Tabela 4.4: Tipo de Negócio

Tipo negócio	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Desvio padrão do custo	Frequência	Rácio Sinistralidade	Prémio Risco
Novo negócio	2014	79 974	38 859	210,2	5 742	7 443 731	1 296	901	14,8%	91,2%	191,6
	2015	56 750	26 944	190,8	3 465	4 501 672	1 299	1 019	12,9%	87,5%	167,1
	2016	51 337	25 666	182,6	3 201	3 616 661	1 130	753	12,5%	77,2%	140,9
	2017	73 717	31 686	181,8	3 654	4 218 668	1 154	888	11,5%	73,3%	133,1
Renovações	2014	185 076	158 378	215,0	17 697	23 673 172	1 338	1 620	11,2%	69,5%	149,5
	2015	199 906	169 636	212,4	19 328	24 249 911	1 255	1 291	11,4%	67,3%	143,0
	2016	190 004	161 922	207,2	17 456	21 088 417	1 208	1 652	10,8%	62,9%	130,2
	2017	181 611	156 380	202,8	15 495	17 859 760	1 153	1 110	9,9%	56,3%	114,2

A análise da tabela 4.4 mostra que o novo negócio apresenta piores indicadores (custo médio, frequência e rácio de sinistralidade) mais elevados e consequentemente prémio de risco mais elevado, do que as renovações. Tal deve-se ao novo negócio apresentar prémios mais competitivos e em linha de conta com o mercado, e ao facto de alguma parte do novo negócio ser um risco maior, uma vez que a companhia não conhece o histórico de sinistralidade do cliente.

- Tipo de Combustível

Tabela 4.5: Tipo de Combustível

Tipo combustível	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Desvio-Padrão Custos	Frequência	Rácio Sinistralidade	Prémio Risco
Gasolina	2014	98 985	72 391	172,7	6 261	8 017 817	1 041	1 281	8,6%	64,1%	110,8
	2015	95 822	72 597	168,0	6 075	7 217 230	815	1 188	8,4%	59,2%	99,4
	2016	86 885	67 319	165,5	5 257	6 743 348	1 761	1 283	7,8%	60,5%	100,2
	2017	87 839	64 579	165,1	4 519	5 391 131	1 130	1 193	7,0%	50,6%	83,5
Gasóleo	2014	151 564	112 274	255,4	16 948	22 855 449	1 708	1 349	15,1%	79,7%	203,6
	2015	149 014	113 619	247,4	16 508	21 296 632	1 481	1 290	14,5%	75,7%	187,4
	2016	143 623	110 599	236,8	15 224	17 706 451	1 386	1 163	13,8%	67,6%	160,1
	2017	157 974	114 997	226,3	14 445	16 481 498	1 033	1 141	12,6%	63,3%	143,3
Outros	2014	14 502	12 573	83,0	230	243 637	235	1 060	1,8%	23,4%	19,4
	2015	11 821	10 363	83,0	210	237 720	322	1 131	2,0%	27,6%	22,9
	2016	10 834	9 670	93,5	176	255 279	475	1 450	1,8%	28,2%	26,4
	2017	9 514	8 490	92,9	185	205 798	281	1 114	2,2%	26,1%	24,2

A tabela 4.5 indica que os veículos a gasóleo apresentam piores indicadores (frequência e rácio de sinistralidade mais elevados) e consequentemente prémio de risco mais elevado, do que os veículos a gasolina e do que os veículos com outros tipos de combustível (elétricos, híbridos e a gás natural). Esta diferença pode dever-se ao facto de que a maioria dos carros das empresas são a gasóleo, o que significa que esta variável pode estar relacionada com a variável tipo de cliente. A fim de comprovar esta hipótese fez-se uma análise combinando as duas variáveis, tipo de cliente e tipo de combustível.

Tabela 4.6: Tipo de Cliente e Tipo de Combustível

Tipo cliente	Tipo combustivel	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Desvio-Padrão Custos	Frequência	Rácio Sinistralidade	Prémio Risco
Empresa	Gasolina	2014	3 224	2 349	275,0	249	514 139	2 068	1 324	10,6%	79,6%	218,9
		2015	2 929	2 219	266,6	224	305 954	1 363	1 039	10,1%	51,7%	137,9
		2016	2 668	2 044	249,9	208	237 940	1 144	611	10,2%	46,6%	116,4
		2017	2 655	1 942	243,2	179	208 680	1 164	510	9,2%	44,2%	107,5
	Gasóleo	2014	28 757	21 496	357,1	4 409	6 200 442	1 406	1 989	20,5%	80,8%	288,4
		2015	26 840	20 536	349,2	4 111	5 623 553	1 368	1 712	20,0%	78,4%	273,8
		2016	26 006	19 729	333,6	3 852	4 409 613	1 145	1 024	19,5%	67,0%	223,5
		2017	29 578	21 172	313,3	3 769	4 331 255	1 149	953	17,8%	65,3%	204,6
	Outros	2014	1 068	870	170,1	36	38 972	1 074	313	4,2%	26,3%	44,8
		2015	959	817	165,6	29	35 932	1 256	345	3,5%	26,6%	44,0
		2016	977	839	175,6	31	46 133	1 498	472	3,7%	31,3%	55,0
		2017	969	809	170,3	37	33 033	883	238	4,6%	24,0%	40,8
Particular	Gasolina	2014	95 761	70 042	169,3	6 013	7 503 679	1 248	1 030	8,6%	63,3%	107,1
		2015	92 893	70 378	164,9	5 851	6 911 275	1 181	807	8,3%	59,6%	98,2
		2016	84 217	65 275	162,9	5 049	6 505 408	1 288	1 785	7,7%	61,2%	99,7
		2017	85 184	62 637	162,7	4 339	5 182 451	1 194	1 144	6,9%	50,8%	82,7
	Gasóleo	2014	122 806	90 777	231,3	12 539	16 655 007	1 328	1 634	13,8%	79,3%	183,5
		2015	122 174	93 083	225,0	12 397	15 673 079	1 264	1 424	13,3%	74,8%	168,4
		2016	117 616	90 870	215,8	11 372	13 296 838	1 169	1 453	12,5%	67,8%	146,3
		2017	128 396	93 825	206,6	10 677	12 150 243	1 138	1 050	11,4%	62,7%	129,5
	Outros	2014	13 434	11 703	76,5	194	204 665	1 057	228	1,7%	22,9%	17,5
		2015	10 861	9 546	75,9	181	201 789	1 112	320	1,9%	27,8%	21,1
		2016	9 857	8 831	85,7	145	209 146	1 440	475	1,6%	27,6%	23,7
		2017	8 545	7 681	84,7	147	172 765	1 172	285	1,9%	26,5%	22,5

Através da análise da tabela 4.6 conclui-se que, efetivamente, os clientes do tipo empresa têm mais carros a gasóleo, mas verifica-se também que os clientes do tipo particular têm mais carros a gasóleo. Contudo, nos particulares a diferença entre carros a gasóleo e a gasolina não é tão expressiva, uma vez que, em 2017, os carros a gasóleo representavam 89% e 58% do total de veículos para clientes do tipo empresa e particular, respetivamente. Isto deve-se ao facto de cerca de 60% da amostra ser constituída por veículos a gasóleo. Pelo que podemos concluir que existe uma relação entre os clientes do tipo empresa e os veículos a gasóleo.

#### 4.4. Análise de tendências e das medidas tomadas no passado

Para compreender o impacto das medidas tomadas pela Companhia no período em estudo, foi feita uma análise de tendências, que nos ajuda a perceber se existem padrões temporais nos dados e foram estudadas as medidas tomadas.

##### 4.4.1. Análise de Tendências

A análise de tendências foi feita recorrendo a um modelo de regressão linear simples, em que foi calculada a variação média anual de cada um dos indicadores chave o que corresponde ao coeficiente  $a$  do modelo  $y=ax+b$ . Deste modo, a variação anual, para cada indicador, é dada por:

$$\text{Variação Média Anual} = \frac{\text{covar}(x,y)}{\text{var}(x)} \quad (4.1)$$

onde  $x$  corresponde à variável ano e  $y$  corresponde ao valor assumido pelo indicador em estudo para cada um dos anos.

Para os indicadores chave mais importantes foi feita a representação gráfica da evolução no período e traçada uma reta de regressão, com indicação da equação do modelo e do coeficiente de determinação,  $R^2$ .

Na tabela 4.7 são apresentados os indicadores chave para o global da carteira e os resultados para a estimativa da variação média anual dos mesmos, denominada por Variação Anual.

Tabela 4.7: Tabela Global com a Variação Anual

Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Frequência	Rácio Sinistralidade	Prémio Risco
2014	265 051	197 237	214,1	23 439	31 116 904	1 328	11,9%	73,7%	157,8
2015	256 656	196 580	209,4	22 793	28 751 582	1 261	11,6%	69,8%	146,3
2016	241 341	187 588	203,8	20 657	24 705 078	1 196	11,0%	64,6%	131,7
2017	255 328	188 066	199,3	19 149	22 078 427	1 153	10,2%	58,9%	117,4
Variação Anual	- 4 448	- 3 650	- 5,0	- 1 501	- 3 116 193	- 59	- 0,006	- 0,050	- 13,6

Com base nos resultados apresentados na tabela 4.7 conclui-se que, em média e por ano, tem havido uma redução de 3 650 apólices em risco, uma redução de 5 euros no prémio médio, uma redução de 1 501 sinistros e consequentemente uma redução de, aproximadamente, 3 milhões de euros em custos com sinistros. Isto significa que houve uma redução de, em média por ano, 0.6 p.p. na frequência e de 5 p.p. no rácio de sinistralidade. No entanto, observa-se que o último ano teve um comportamento ligeiramente diferente dos restantes, uma vez que há pela primeira vez um crescimento do número de apólices, continuando-se a verificar uma redução do custo médio e da frequência, o que produz um rácio de sinistralidade mais baixo.

Seguidamente são apresentados os gráficos da análise de tendências para a Exposição, Prémio Médio, Número de Sinistros, Custo Médio, Frequência de Sinistralidade e Rácio de Sinistralidade.

- Exposição

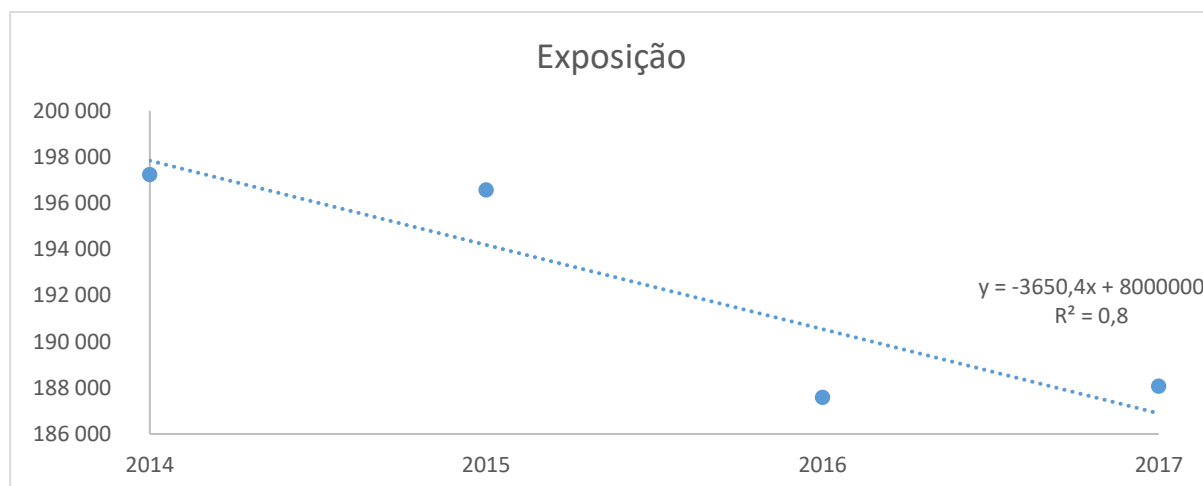


Gráfico 4.1: Exposição

Através da análise do gráfico 4.1 conclui-se que a variação da exposição não foi constante ao longo do tempo, uma vez que houve uma ligeira diminuição entre 2014 e 2015 (-657), entre 2015 e 2016 houve uma diminuição mais acentuada (-8 992) e entre 2016 e 2017 já houve um ligeiro aumento (478). Mas ainda assim, conclui-se que a variação foi aproximadamente linear, uma vez que o coeficiente de determinação está próximo de 1 ( $R^2=0.8$ ).

- Prémio Médio

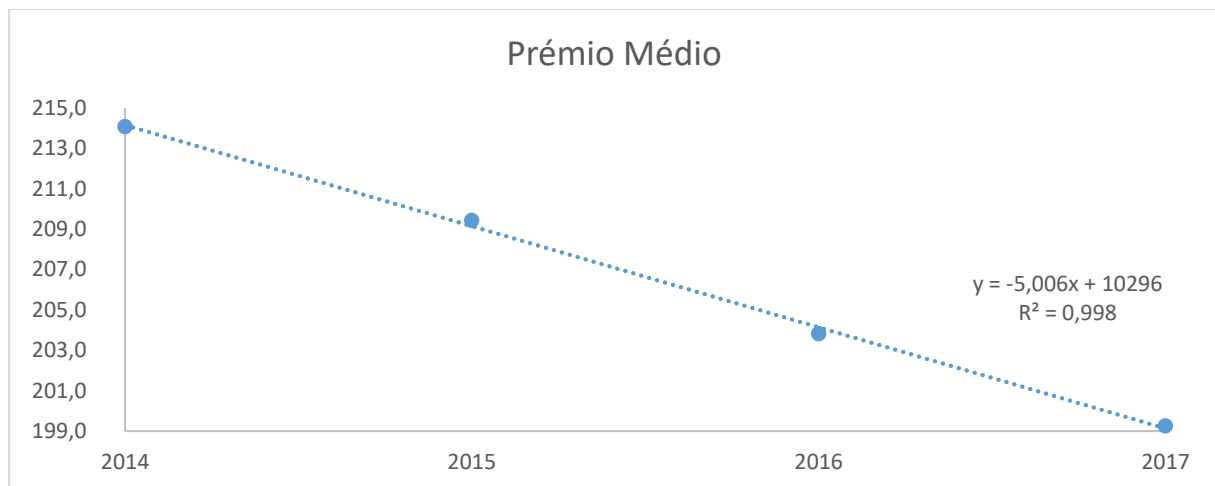


Gráfico 4.2: Prémio Médio

O gráfico 4.2 permite concluir que a variação do prémio médio foi praticamente linear, uma vez que houve uma diminuição constante ao longo dos quatro anos de aproximadamente 5 euros em cada um dos anos e uma vez que o coeficiente de determinação é aproximadamente 1. O que é normal, uma vez que os prémios apesar de serem controlados pela Companhia variam consoante a composição da carteira. Uma vez que a composição da carteira terá sido fortemente controlada e melhorada durante o período em análise, a redução do prémio tem o efeito da redução do número de apólices em risco.

- Número de Sinistros

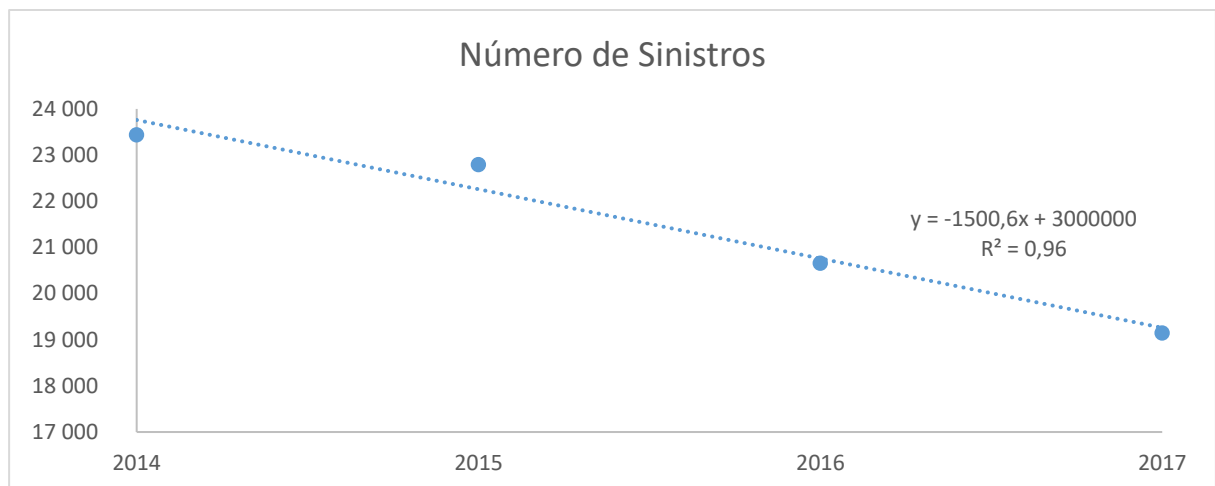


Gráfico 4.3: Número de Sinistros

O gráfico 4.3 mostra que a variação do número de sinistros foi linear, uma vez que houve uma constante diminuição, apesar de entre 2014 e 2015 a diminuição (-646) ter sido inferior à diminuição ocorrida nos outros períodos (aproximadamente 2 mil). Podemos afirmar que é linear uma vez que temos um coeficiente de determinação muito próximo de 1.

- Custo Médio

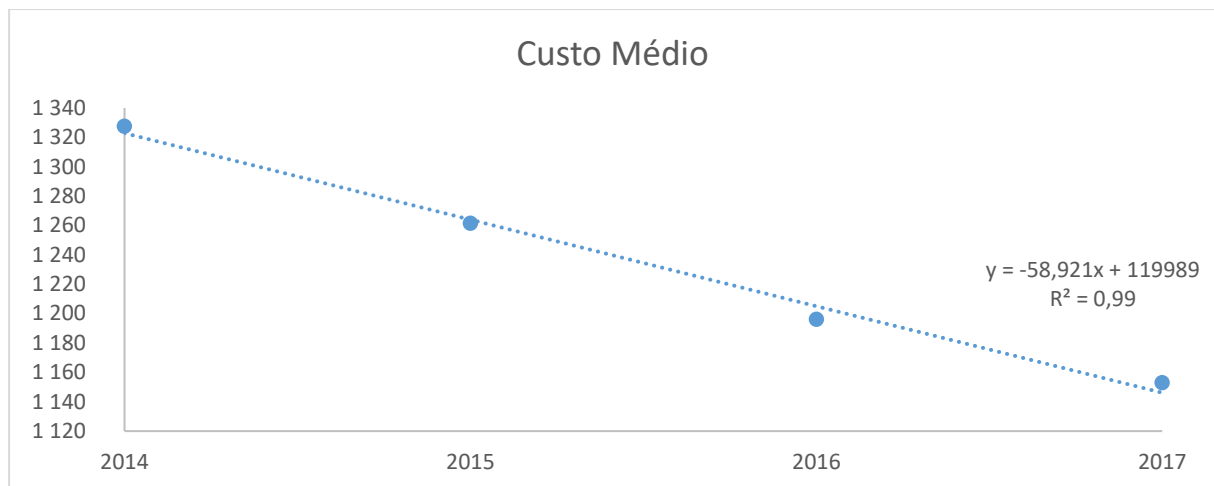


Gráfico 4.4: Custo Médio

Através da análise do gráfico 4.4 conclui-se que a variação do custo médio foi linear, uma vez que houve uma constante diminuição e que o coeficiente de determinação é aproximadamente 1, apesar de entre 2016 e 2017 a diminuição (-40) ter sido ligeiramente inferior à variação nos restantes períodos (-66).

- Frequência de Sinistralidade

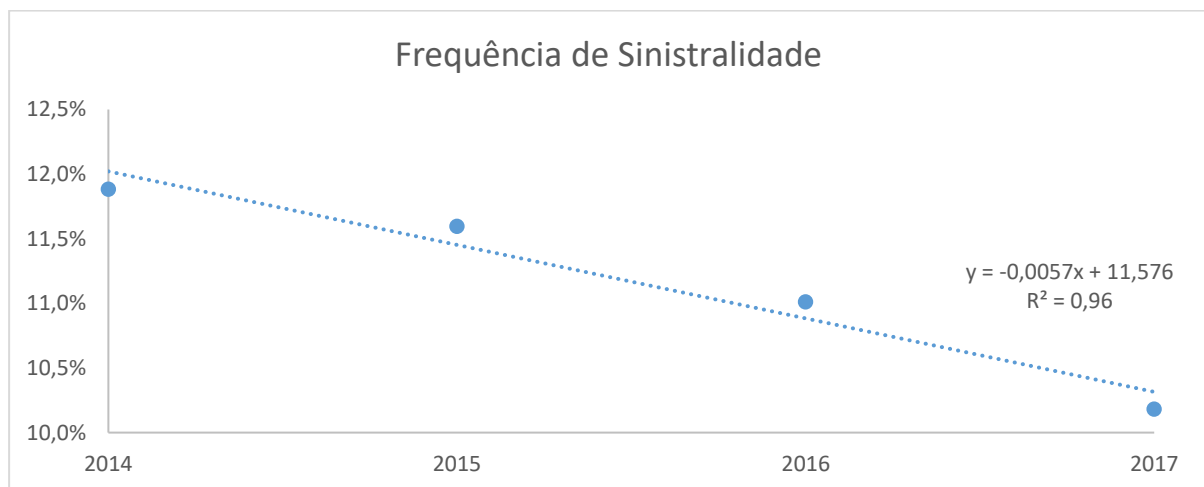


Gráfico 4.5: Frequência de Sinistralidade

A partir do gráfico 4.5 conclui-se que a variação da frequência foi linear, uma vez que a diminuição foi constante durante os 4 anos e o coeficiente de determinação é próximo de 1. No entanto, apesar de entre 2014 e 2015 a diminuição ter sido de apenas 0.3 p.p., entre 2015 e 2016 houve uma diminuição de 0.6 p.p. e entre 2016 e 2017 a diminuição atingiu os 0.8 p.p.



- Rácio de Sinistralidade

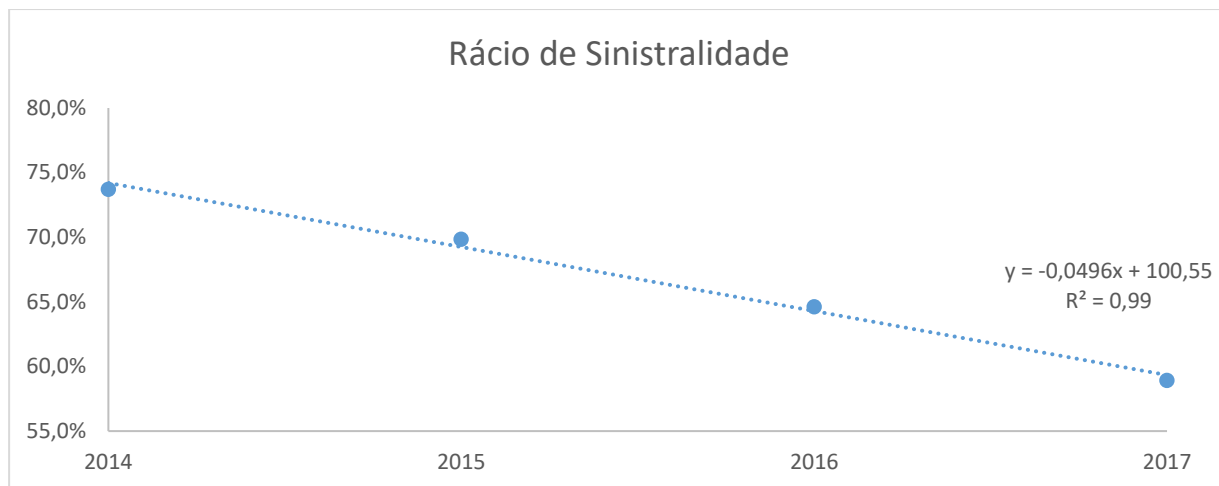


Gráfico 4.6: Rácio de Sinistralidade

Finalmente, da análise do gráfico 4.6 conclui-se que a variação do rácio de sinistralidade foi linear, uma vez que a diminuição foi sempre constante e o coeficiente de determinação é aproximadamente igual a 1, apesar da diminuição entre os anos 2014 e 2015 ter sido ligeiramente inferior (-4 p.p.) aos restantes períodos (aproximadamente 5,5 p.p.).

Conclui-se assim que entre os anos de 2014 e 2016 houve uma redução do número de apólices em risco acompanhada de diminuição da Frequência de Sinistralidade e do custo médio dos sinistros, e consequentemente uma diminuição do rácio de sinistralidade. Pode-se também concluir que apesar de em 2017 o número de apólices em risco ter aumentado ligeiramente, continuamos a baixar a frequência e o rácio de sinistralidade.

#### 4.4.2. Análise das medidas do passado

A maior diminuição no número de apólices em risco, no prémio médio e no número de sinistros ter ocorrido entre 2015 e 2016 deveu-se ao facto de em 2015 terem sido tomadas algumas medidas, entre as quais:

- Medida 1: Alteração do nível de aceitação de seguros fracionados por rating de mediador, de forma a reduzir o número de apólices fracionadas (Aplica-se a novas apólices).
- Medida 2: Alteração da tarifa de danos próprios, sofrendo um aumento de 10%. (Aplica-se a novas apólices e a apólices renovadas)
- Medida 3: Alteração do nível de aceitação por franquia e por idade do veículo em veículos com danos próprios, de forma a limitar o número de carros mais antigos com seguro de danos próprios (Aplica-se a novas apólices).
- Medida 4: Definição de novas regras de saneamento, de forma a sanear clientes menos rentáveis e com piores resultados, isto é, apólices com um ou mais sinistros no último ano e rácio de sinistralidade superior a 70% (Aplica-se à carteira renovada).

Foi também medido o impacto destas medidas mais ao detalhe, usando para isso algumas tabelas que foram construídas inicialmente para as variáveis referidas no início do capítulo e que se enquadram em cada uma das medidas. Será considerada a mesma ordem das medidas apresentada em cima.

A medida 1, consistiu em limitar a venda de apólices fracionadas a agentes com *rating* "prata" das Grandes Cidades, Litoral e Ilhas e agentes "bronze", exceto para clientes de alto valor (clientes com mais do que uma apólice na Companhia). Para tal foi feita uma análise ao tipo de pagamento fracionado por *rating* do agente (em vigor em 2015), tal como é apresentado na tabela 4.8. Na tabela 4.9 é apresentado o impacto da medida.

Tabela 4.8: Tipo de Pagamento Fracionado por Rating de Agente

Rating Agente	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Frequência	Rácio Sinistralidade	Prémio Risco
Ouro	2014	33 203	24 592	233,8	3 550	4 698 383	1 324	14,4%	81,7%	191,1
	2015	33 160	24 923	229,0	3 679	4 613 708	1 254	14,8%	80,8%	185,1
	2016	30 455	24 169	222,6	3 439	3 825 994	1 113	14,2%	71,1%	158,3
	2017	22 934	17 223	220,3	2 138	2 266 350	1 060	12,4%	59,7%	131,6
Prata	2014	46 249	32 303	227,3	5 267	7 263 698	1 379	16,3%	98,9%	224,9
	2015	40 748	31 292	222,3	4 908	6 392 653	1 302	15,7%	91,9%	204,3
	2016	31 543	24 339	223,4	3 450	4 086 077	1 185	14,2%	75,2%	167,9
	2017	18 821	14 340	233,8	1 875	2 231 731	1 190	13,1%	66,6%	155,6
Bronze	2014	6 131	4 008	232,5	818	1 379 970	1 686	20,4%	148,1%	344,3
	2015	4 395	3 359	236,1	634	874 217	1 380	18,9%	110,3%	260,3
	2016	2 087	1 520	250,9	247	361 568	1 461	16,3%	94,8%	237,9
	2017	2 080	1 631	234,3	220	319 640	1 453	13,5%	83,7%	196,0

Tabela 4.9: Impacto da redução da venda de apólices fracionadas

	Variação da exposição 2015-2016	Variação do nºsinistros 2015-2016	Variação do custo médio 2015-2016	Variação da frequência 2015-2016	Variação do Rácio de Sinistralidade 2015-2016
Ouro	-3 %	-7 %	-11 %	-4 p.p.	-12 p.p.
Prata	-22 %	-30 %	-9 %	-10 p.p.	-18 p.p.
Bronze	-55 %	-61 %	6 %	-14 p.p.	-14 p.p.

Através da análise da tabela 4.9, podemos concluir que esta medida levou, tal como já era expectável, a uma grande redução da exposição de apólices fracionadas nos agentes bronze, reduzimos neste segmento, mais de metade das apólices (-55%), o que levou a uma redução de 61% do número de sinistros. Apesar de termos reduzido significativamente o número de sinistros, em 2015 obtivemos um custo médio mais elevado, uma vez que houve um aumento de 6% nos custos. Conseguimos em todo o caso reduzir em 14 p.p. a frequência de sinistralidade e o rácio de sinistralidade neste segmento.

No segmento dos agentes prata tivemos uma redução de 22% no número de apólices, acompanhada de uma redução de 30% dos sinistros, com custo médio ligeiramente mais baixo (-9%). Tivemos assim para os agentes prata uma redução de 10 p.p. da frequência e uma redução de 18 p.p. do rácio de sinistralidade.

Para os agentes ouro conseguimos também uma ligeira redução do número de apólices (-3%) e do número de sinistros (-7%), o que é expectável visto que estes são o melhor segmento do *rating* de agente. Ainda assim conseguimos neste segmento reduzir 12 p.p. o rácio de sinistralidade, pela acentuada redução do custo médio (-11%).

Para a *medida 2*, que consistiu num aumento da tarifa de danos próprios em 10%, foi feito um estudo às apólices com esta cobertura, tal como podemos ver na tabela 4.10. A tabela 4.11 mostra o impacto desta medida, associado ao impacto das outras medidas apresentadas.

Tabela 4.10: Apólices com Danos Próprios (DP)

Ano	Nº Apólices	Exposição	Nº sinistros	Custo sinistros	Custo médio	Frequência	Rácio Sinistralidade	Prémio Risco
2014	42 235	33 257	6 978	13 471 963	1 931	21,0%	89,1%	405,1
2015	38 684	30 610	6 394	11 575 042	1 810	20,9%	82,0%	378,2
2016	32 233	25 826	5 053	8 112 913	1 605	19,6%	68,0%	314,1
2017	30 184	23 694	4 291	6 785 985	1 581	18,1%	62,9%	286,4

Tabela 4.11: Impacto do aumento da tarifa de DP em 10%

	Variação da exposição 2015-2016	Variação do nº sinistros 2015-2016	Variação do custo médio 2015-2016	Variação da frequência 2015-2016	Variação do Rácio de Sinistralidade 2015-2016
Total	-16 %	-21 %	-11 %	-6 p.p.	-17 p.p.

A análise da tabela 4.11 permite concluir que o aumento da tarifa de DP levou a uma redução de 16% das apólices com danos próprios, acompanhado de uma redução de 21% do número de sinistros, e de um custo médio mais baixo (-11%). Conseguimos assim reduzir o rácio de sinistralidade em 17 p.p. e a frequência em 6 p.p..

A *medida 3* consistiu em baixar a idade limite dos veículos de 4 para 2 anos, em novas apólices com danos próprios e sem franquia. Para novas apólices com danos próprios e franquia de 2% a idade passou de 10 para 6 anos. Assim, para esta medida foi feito um estudo por idade do veículo para apólices com danos próprios e franquias de 0% e 2%. Na tabela 4.12 são apresentadas as apólices sem franquia e na tabela 4.14 as apólices com franquia de 2%. Nas tabelas 4.13 e 4.15 são apresentados os impactos, respetivamente, para apólices sem franquia e com franquia de 2%.

Tabela 4.12: Apólices com DP Sem Franquia

Idade Veículo	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Frequência	Rácio Sinistralidade	Prémio Risco
0-2	2014	4 094	3 260	538,4	876	1 538 876	1 758	26,9%	87,7%	472,1
	2015	4 064	3 488	537,5	888	1 924 912	2 168	25,5%	102,7%	551,9
	2016	3 614	2 986	508,1	688	1 035 608	1 506	23,0%	68,3%	346,8
	2017	3 684	2 963	464,5	622	920 597	1 481	21,0%	66,9%	310,7
3-4	2014	4 187	3 674	499,2	933	1 681 984	1 803	25,4%	91,7%	457,8
	2015	2 795	2 504	508,9	649	1 178 697	1 816	25,9%	92,5%	470,7
	2016	2 044	1 871	544,8	521	861 889	1 653	27,9%	84,6%	460,6
	2017	1 944	1 819	547,1	452	701 038	1 551	24,9%	70,5%	385,5

Tabela 4.13: Impacto da medida em apólices com DP sem franquia

	Varição da exposição 2015-2016	Varição do nºsinistros 2015-2016	Varição do custo médio 2015-2016	Varição da frequência 2015-2016	Varição do Rácio de Sinistralidade 2015-2016
0-2	-14 %	-23 %	-31 %	-10 p.p.	-34 p.p.
3-4	-25 %	-20 %	-9 %	8 p.p.	-9 p.p.

Com a análise da tabela 4.13 concluiu-se que a redução da idade dos veículos em apólices com danos próprios sem franquia representou uma diminuição do número de apólices em 25% em veículos com idades entre os 3 e os 4 anos, o que levou a uma diminuição de 20% do número de sinistros e a um custo médio mais baixo (-9%). Conseguimos assim reduzir o rácio de sinistralidade em 9 p.p. neste segmento embora a frequência tenha aumentado ligeiramente (8 p.p.). Diminuímos também o número de apólices de veículos com idade até 2 anos, o que deve estar ligado ao aumento da tarifa de DP e uma diminuição de venda de apólices com pagamento fracionado analisado nas medidas anteriores. Nesse segmento conseguimos uma redução de 34 p.p. do rácio de sinistralidade e de 10 p.p. na frequência.

Tabela 4.14: Apólices com DP Franquia 2%

Idade Veículo	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Frequência	Rácio Sinistralidade	Prémio Risco
0-2	2014	3 298	2 702	507,2	612	1 031 362	1 686	22,6%	75,3%	381,7
	2015	2 961	2 438	520,2	490	1 112 277	2 272	20,1%	87,7%	456,1
	2016	2 775	2 222	500,6	450	726 167	1 614	20,2%	65,3%	326,8
	2017	3 540	2 482	470,3	482	795 713	1 652	19,4%	68,2%	320,6
3-4	2014	5 955	5 064	466,6	1 047	2 179 464	2 081	20,7%	92,2%	430,3
	2015	3 391	2 782	495,2	584	1 199 036	2 053	21,0%	87,0%	431,0
	2016	2 445	2 061	499,2	408	785 091	1 924	19,8%	76,3%	380,8
	2017	2 759	2 187	489,4	442	761 180	1 721	20,2%	71,1%	348,1
5-6	2014	5 759	4 936	465,4	1 009	2 096 171	2 078	20,4%	91,3%	424,7
	2015	5 256	4 431	467,1	902	1 765 804	1 958	20,4%	85,3%	398,5
	2016	4 182	3 577	471,9	706	1 216 303	1 722	19,7%	72,1%	340,1
	2017	2 746	2 320	462,9	420	753 437	1 793	18,1%	70,2%	324,8
7-10	2014	5 385	4 712	474,5	972	2 270 611	2 335	20,6%	101,5%	481,9
	2015	5 678	4 788	484,7	1 021	1 749 213	1 714	21,3%	75,4%	365,3
	2016	4 019	3 594	496,7	648	1 064 101	1 642	18,0%	59,6%	296,1
	2017	3 776	3 415	486,3	613	1 065 209	1 739	17,9%	64,1%	311,9

Tabela 4.15: Impacto da medida em apólices com DP com franquia 2%

	Varição da exposição 2015-2016	Varição do nºsinistros 2015-2016	Varição do custo médio 2015-2016	Varição da frequência 2015-2016	Varição do Rácio de Sinistralidade 2015-2016
0-2	-9 %	-8 %	-29 %	1 p.p.	-26 p.p.
3-4	-26 %	-30 %	-6 %	-6 p.p.	-12 p.p.
5-6	-19 %	-22 %	-12 %	-3 p.p.	-16 p.p.
7-10	-25 %	-37 %	-4 %	-15 p.p.	-21 p.p.

Da análise da tabela 4.15 conclui-se que, com a redução da idade dos veículos em apólices com danos próprios e com franquia de 2%, obteve-se uma redução de 25% das apólices em veículos com idades entre 7 e 10 anos, o que significou uma redução de 37% no número de sinistros e um custo médio

ligeiramente mais baixo (-4%). Esta medida também levou a uma redução de 21 p.p. do rácio de sinistralidade e de 15 p.p. na frequência. Concluímos também que, em veículos com idade entre os 3 e os 4 anos, houve uma redução de 26% do número de apólices, o que representou uma redução de 30% do número de sinistros e uma redução de 12 p.p. do rácio de sinistralidade. No segmento dos veículos com idade até 2 anos vemos que, apesar de termos tido uma ligeira redução do número de apólices (-9%), obteve-se uma redução de 26 p.p. do rácio de sinistralidade, pois houve uma redução do custo médio substancial (-29%).

Na *medida 4* foi implementado um processo que através de um sistema de pontuação identifica e elimina apólices com uma elevada frequência e um elevado rácio de sinistralidade. Contribuem para este processo de pontos variáveis como o número de sinistros durante o último ano, a antiguidade da apólice na carteira, o *rating* do cliente, o *rating* do agente, a forma de pagamento e se a apólice tem a cobertura de Danos Próprios. Qualquer apólice com um ou mais sinistros durante o último ano e com um rácio de Sinistralidade superior a 70% pode ser anulada. Assim, para esta medida foi feito um estudo ao número de apólices saneadas, como se pode ver na tabela 4.16, e respetivo impacto na carteira, apresentado na tabela 4.17. Na tabela 4.17, a variação do número de apólices saneadas é calculada com base na tabela 4.16 (apólices saneadas), já a variação da frequência e do rácio de sinistralidade são calculados com base na tabela 4.8 (total da carteira).

Tabela 4.16: Apólices Saneadas

Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Custo Médio	Frequência	Rácio Sinistralidade	Prémio Risco
2014	2 299	1 679	304,0	1 381	2 124 495	1 539	82,2%	416,3%	1 265,4
2015	3 776	3 132	284,1	3 344	5 461 163	1 633	106,8%	613,7%	1 743,8
2016	4 904	3 511	265,3	3 505	5 800 982	1 655	99,8%	622,8%	1 652,0
2017	2 576	1 523	251,2	1 465	2 408 084	1 644	96,2%	629,2%	1 580,9

Tabela 4.17: Impacto das apólices saneadas

	Variação do nºapólices saneadas 2015-2016	Variação da frequência 2015-2016	Variação do Rácio de Sinistralidade 2015-2016
Total	30 %	-5 p.p.	-7 p.p.

Com a análise da tabela 4.17 conclui-se que, em 2016, foram saneadas mais 30% das apólices, o que representou, no total da carteira, uma diminuição de 5 p.p. na frequência e uma diminuição de 7 p.p. no rácio de sinistralidade.

## 4.5. Companhia em Análise vs. Mercado Segurador

Foi feita uma análise comparativa entre a Companhia e o mercado segurador de forma a compreender qual a quota de mercado total da Companhia e a quota por distrito. Esta análise foi feita para os anos de 2016 e 2017, para ser possível observar as variações.

As tabelas 4.18 e 4.19 apresentam o número de veículos em carteira na Companhia, o número de veículos no mercado segurador, a quota de mercado da Companhia, que corresponde ao quociente entre o número de veículos da Companhia e o número de veículos do mercado, e o rácio de sinistralidade da Companhia por distrito para os anos 2016 e 2017, respetivamente.

Tabela 4.18: Quota de Mercado por Distrito 2016

Distrito	Nº Veículos Companhia	Nº Veículos Mercado	Quota de Mercado Companhia	Rácio de Sinistralidade Companhia
Açores	1 576	141 132	1,1%	64,3%
Aveiro	17 365	532 166	3,3%	54,9%
Beja	3 666	107 968	3,4%	50,6%
Braga	28 227	585 419	4,8%	64,8%
Bragança	2 831	115 420	2,5%	54,5%
Castelo Branco	9 679	142 870	6,8%	69,1%
Coimbra	15 253	330 153	4,6%	71,9%
Évora	6 448	115 123	5,6%	51,7%
Faro	11 539	339 516	3,4%	55,9%
Guarda	3 249	134 118	2,4%	46,7%
Leiria	14 702	397 748	3,7%	58,5%
Lisboa	18 824	1 438 144	1,3%	58,7%
Madeira	4 814	146 873	3,3%	76,2%
Portalegre	4 814	80 359	6,0%	56,5%
Porto	48 650	1 048 070	4,6%	70,6%
Santarém	8 270	334 130	2,5%	56,8%
Setúbal	8 120	472 995	1,7%	103,8%
Viana do Castelo	14 069	185 028	7,6%	68,5%
Vila Real	5 633	152 208	3,7%	63,2%
Viseu	13 516	293 135	4,6%	52,7%

Total	241 244	7 092 575	3,4%	64,6%
-------	---------	-----------	------	-------

Tabela 4.19: Quota de Mercado por Distrito 2017

Distrito	Nº Veículos Companhia	Nº Veículos Mercado	Quota de Mercado Companhia	Rácio de Sinistralidade Companhia
Açores	1 730	146 940	1,2%	47,5%
Aveiro	18 246	558 699	3,3%	55,9%
Beja	4 021	114 553	3,5%	49,5%
Braga	28 508	608 102	4,7%	55,0%
Bragança	3 737	119 389	3,1%	47,1%
Castelo Branco	11 120	148 703	7,5%	57,6%
Coimbra	14 057	340 170	4,1%	57,4%
Évora	7 550	117 996	6,4%	63,7%
Faro	15 017	352 087	4,3%	64,6%
Guarda	4 079	138 625	2,9%	44,7%
Leiria	14 687	405 781	3,6%	60,4%
Lisboa	18 211	1 487 093	1,2%	57,6%
Madeira	4 465	148 443	3,0%	49,0%
Portalegre	5 128	82 904	6,2%	49,3%
Porto	51 579	1 085 822	4,8%	64,1%
Santarém	6 989	346 010	2,0%	72,5%
Setúbal	8 909	486 270	1,8%	54,4%
Viana do Castelo	15 502	191 640	8,1%	55,9%
Vila Real	7 029	158 389	4,4%	76,7%
Viseu	14 663	304 642	4,8%	51,1%

Total	255 226	7 342 258	3,5%	58,9%
-------	---------	-----------	------	-------

Na tabela 4.20 é apresentada a taxa de variação do número de veículos da Companhia, do número de veículos do mercado, da quota de mercado da Companhia e do rácio de sinistralidade da Companhia entre os anos 2016 e 2017 por distrito.

Tabela 4.20: Variação por Distrito

Distrito	Variação NºVeículos Companhia 16-17	Variação NºVeículos Mercado 16-17	Variação Quota de Mercado Companhia 16-17	Variação Rácio Sinistralidade Companhia 16-17
Açores	9,8%	4,1%	5,4%	-26,2%
Aveiro	5,1%	5,0%	0,1%	1,9%
Beja	9,7%	6,1%	3,4%	-2,3%
Braga	1,0%	3,9%	-2,8%	-15,2%
Bragança	32,0%	3,4%	27,6%	-13,6%
Castelo Branco	14,9%	4,1%	10,4%	-16,7%
Coimbra	-7,8%	3,0%	-10,6%	-20,2%
Évora	17,1%	2,5%	14,2%	23,2%
Faro	30,1%	3,7%	25,5%	15,7%
Guarda	25,5%	3,4%	21,4%	-4,2%
Leiria	-0,1%	2,0%	-2,1%	3,1%
Lisboa	-3,3%	3,4%	-6,4%	-1,8%
Madeira	-7,2%	1,1%	-8,2%	-35,7%
Portalegre	6,5%	3,2%	3,3%	-12,7%
Porto	6,0%	3,6%	2,3%	-9,2%
Santarém	-15,5%	3,6%	-18,4%	27,7%
Setúbal	9,7%	2,8%	6,7%	-47,7%
Viana do Castelo	10,2%	3,6%	6,4%	-18,4%
Vila Real	24,8%	4,1%	19,9%	21,4%
Viseu	8,5%	3,9%	4,4%	-3,2%
Total	5,8%	3,5%	2,2%	-8,8%

Através da análise da tabela 4.20, podemos comparar a variação do número de veículos da Companhia com a variação do mercado e ver qual o impacto dessa variação no resultado da Companhia, através da variação do rácio de sinistralidade.

Pode-se ver que em Bragança o crescimento da Companhia (32%) foi bastante superior ao crescimento do mercado (3,4%), e o mesmo aconteceu em Faro, em que a Companhia cresceu 30,1% e o mercado apenas cresceu 3,7%. Reparamos também que em Santarém o mercado (3,6%) cresceu mais que a Companhia (-15,5%).

Estes valores estão de acordo com a evolução da quota de mercado, onde concluímos também que em certos distritos o aumento da quota foi significativo, como foi o caso de Bragança (27,6%) e Faro (25,5%) e que em Santarém (-18,4%) houve uma grande perda de quota de mercado.

No entanto, um crescimento da quota de mercado nem sempre melhora o resultado da Companhia, como acontece, por exemplo, em Faro em que o crescimento da quota de mercado faz com que o rácio de sinistralidade aumente 15,7%.

Conclui-se também que a quota global da Companhia aumentou ligeiramente (2,2%), acompanhado de uma redução do rácio de sinistralidade da Companhia em 8,8%.

## 5. Métodos de gestão da carteira

Neste capítulo vamos descrever as análises que propomos com o objetivo de facilitar a identificação dos segmentos onde devemos agir, isto é, os métodos que pensamos que podem ajudar o gestor da carteira automóvel a rapidamente identificar quais os níveis e de que variáveis que estão a causar problemas.

### 5.1. Tabela Resumo

Para identificar rapidamente onde agir, foram construídas duas tabelas em Excel, que tiveram como ponto de partida a junção de todas as variáveis que consideramos mais importantes (anos de carta, antiguidade da apólice, distrito, idade do condutor, idade do veículo, marca, tipo de cliente, tipo de combustível, tipo de negócio, tipo de pagamento, tipo de veículo e zona geográfica) com os respetivos indicadores chave, já indicados no capítulo anterior, para os anos entre 2014 e 2017. Foi construída uma tabela para uma análise global da carteira e outra só considerando a cobertura de Responsabilidade Civil. Na tabela com a análise global da carteira, foram também acrescentados os mesmos indicadores para o mercado segurador automóvel no mesmo período. Foi adicionalmente calculado o percentil 95 dos custos, isto é, o valor abaixo do qual se encontram 95% do custo de todos os sinistros.

De seguida, foi feita uma alteração, através de uma condição em Excel (função SE), nos indicadores chave Frequência de Sinistralidade e Rácio de Sinistralidade, de forma a só serem visíveis os níveis de variáveis cujos indicadores se encontram acima dos limites desejáveis e que podem ter de ser alvo de ação no futuro. Foi usado como limite para o rácio de sinistralidade o valor de 57%, valor que foi definido pela Companhia, e para a frequência foi usado um limite de 10%, valor correspondente à frequência global em 2017.

Foi também construída uma função em Excel que, através do limite superior de um intervalo de confiança para o valor médio da empresa, nos diz se o custo médio do segmento é inferior a esse limite ou se excedeu o mesmo.

O programa permite que seja o utilizador final quem define quais os limites que quer usar para o rácio de sinistralidade e para a frequência e permite ainda que o utilizador defina qual o nível de confiança que quer usar no intervalo de confiança para o valor médio. Para tal é apresentada ao utilizador final uma tabela como a tabela 5.1.

Tabela 5.1: Tabela para definir os limites utilizados

Límites	Valor
Rácio Sinistralidade	0,57
Frequência	0,10
Nível de confiança	0,95

Para a construção do intervalo de confiança foi considerado um nível de confiança,  $\alpha$ , de 0,1. A tabela 5.2 apresenta as médias (custo médio global por sinistro),  $\bar{x}$ , os desvios padrão das amostras (desvio padrão global dos custos),  $s$ , as dimensões das amostras (número de sinistros total),  $n$ , os quantis da  $t$ -student observados,  $t_{n-1;1-\alpha}$  considerados para a construção do intervalo e o valor obtido para o limite superior do intervalo.



Tabela 5.2: Resumo dos resultados para o cálculo dos intervalos de confiança para custo médio

Ano	$\bar{x}$	s	n	$t_{n-1;1-\alpha}$	Limite Superior
2014	1 327,58	1 441	23 439	1,65	1 343,07
2015	1 261,42	1 236	22 793	1,65	1 274,88
2016	1 195,97	1 507	20 657	1,65	1 213,22
2017	1 152,99	1 051	19 149	1,65	1 165,48

Na tabela 5.3 encontra-se parte da tabela referente aos dados da análise global para 2017. Por uma questão de espaço, este exemplo mostra apenas as variáveis Distrito e Marca e ainda os valores de mercado.

Tabela 5.3: Exemplo tabela resumo para 2017

Variável	Níveis	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Percentil 95 dos custos	Custo Médio	Desvio padrão do custo	Frequência	Rácio Sinistralidade	Prémio Risco	Custo médio
dados mercado	Mercado	2017	6 738 653	-	219,95	871 884	1 105 964 305	-	1 268	-	12,2%	76,1%	154,32	Abaixo do mercado
distrito	Aveiro	2017	18 246	13 448	199,28	1 312	1 498 560	385	1 142	811			111,44	Normal
distrito	Açores	2017	1 730	1 143	145,65	66	79 062	0	1 198	312			69,14	Excedeu
distrito	Beja	2017	4 021	2 835	161,48	234	226 583	255	967	628			79,93	Normal
distrito	Braga	2017	28 508	21 619	213,41	2 219	2 537 836	533	1 144	846	10,3%		117,39	Normal
distrito	Bragança	2017	3 737	2 724	146,78	203	188 132	207	924	458			69,07	Normal
distrito	Castelo Branco	2017	11 120	8 016	185,67	876	856 711	385	978	660	10,9%	57,6%	106,88	Normal
distrito	Coimbra	2017	14 057	10 730	189,82	1 168	1 169 416	421	1 001	638	10,9%	57,4%	108,99	Normal
distrito	Faro	2017	15 017	9 968	167,85	851	1 081 233	305	1 270	627		64,6%	108,47	Excedeu
distrito	Guarda	2017	4 079	2 925	155,76	220	203 805	111	926	415			69,68	Normal
distrito	Leiria	2017	14 687	11 278	200,29	1 189	1 363 816	469	1 147	785	10,5%	60,4%	120,93	Normal
distrito	Lisboa	2017	18 211	13 517	265,04	1 503	2 064 087	770	1 374	891	11,1%	57,6%	152,70	Excedeu
distrito	Madeira	2017	4 465	3 289	150,74	229	243 083	249	1 062	380			73,91	Normal
distrito	Portalegre	2017	5 128	3 823	160,92	332	303 477	284	914	489			79,38	Normal
distrito	Porto	2017	51 579	38 205	223,23	4 588	5 467 192	683	1 192	1 529	12,0%	64,1%	143,10	Excedeu
distrito	Santarém	2017	6 989	5 248	189,92	524	722 637	385	1 380	1 273		72,5%	137,70	Excedeu
distrito	Setúbal	2017	8 909	6 111	206,46	638	685 814	410	1 075	551	10,4%		112,23	Normal
distrito	Viana do Castelo	2017	15 502	11 891	182,99	1 190	1 216 079	393	1 022	636			102,27	Normal
distrito	Vila Real	2017	7 029	4 990	175,56	479	671 804	348	1 404	1 718		76,7%	134,62	Excedeu
distrito	Viseu	2017	14 663	10 872	162,48	865	902 067	242	1 043	764			82,97	Normal
distrito	Évora	2017	7 550	5 360	173,04	458	590 382	286	1 290	2 125		63,7%	110,15	Excedeu
marca	AUDI	2017	7 723	5 745	248,77	804	1 019 132	740	1 267	1 102	14,0%	71,3%	177,41	Excedeu
marca	BMW	2017	9 088	6 811	272,54	1 014	1 206 600	855	1 190	1 135	14,9%	65,0%	177,15	Excedeu
marca	CITROEN	2017	15 281	11 093	209,12	1 292	1 539 714	554	1 191	821	11,7%	66,4%	138,79	Excedeu
marca	FIAT	2017	12 161	8 553	193,63	854	906 496	385	1 062	490			105,99	Normal
marca	FORD	2017	15 653	11 173	204,38	1 317	1 420 283	619	1 079	651	11,8%	62,2%	127,12	Normal
marca	MERCEDES BENZ	2017	13 529	10 373	255,21	1 352	1 583 198	724	1 171	939	13,0%	59,8%	152,63	Excedeu
marca	OPEL	2017	21 652	15 577	189,94	1 589	1 691 648	398	1 064	577	10,2%		108,60	Normal
marca	PEUGEOT	2017	92 021	69 389	173,86	5 437	6 784 900	286	1 248	1 423			97,78	Excedeu
marca	RENAULT	2017	18 847	13 707	212,75	1 496	1 585 724	505	1 060	824	10,9%		115,69	Normal
marca	SEAT	2017	27 476	19 591	207,46	2 336	2 547 594	505	1 090	686	11,9%	62,7%	130,04	Normal
marca	TOYOTA	2017	8 985	6 380	203,09	714	686 346	385	961	585	11,2%		107,58	Normal
marca	Outras	2017	12 912	9 674	204,78	943	1 106 791	433	1 174	714			114,41	Excedeu

Na tabela 5.4 encontra-se uma parte da tabela da análise global, onde são apresentados os piores segmentos em 2017, ou seja, foram filtrados apenas os valores das frequências ou dos rácios de sinistralidade que se encontram em branco e ainda os custos médios que excedem o limite superior para o custo médio obtido através do intervalo de confiança.

Tabela 5.4: Exemplo tabela resumo para os ‘piores’ segmentos em 2017

Variável	Níveis	Ano	Nº Apólices	Exposição	Prémio Médio	Nº Sinistros	Custo Sinistros	Percentil 95 dos custos	Custo Médio	Desvio padrão do custo	Frequência	Rácio Sinistralidade	Prémio Risco	Custo médio
zona geográfica	Grandes Cidade,Litoral, Ilhas	2017	180 057	134 136	211,68	14 325	16 837 344	510	1 175	1 050	10,7%	59,3%	125,52	Excedeu
	Lisboa	2017	18 211	13 517	265,04	1 503	2 064 087	770	1 374	891	11,1%	57,6%	152,70	Excedeu
distrito	Porto	2017	51 579	38 205	223,23	4 588	5 467 192	683	1 192	1 529	12,0%	64,1%	143,10	Excedeu
marca	AUDI	2017	7 723	5 745	248,77	804	1 019 132	740	1 267	1 102	14,0%	71,3%	177,41	Excedeu
marca	BMW	2017	9 088	6 811	272,54	1 014	1 206 600	855	1 190	1 135	14,9%	65,0%	177,15	Excedeu
marca	CITROEN	2017	15 281	11 093	209,12	1 292	1 539 714	554	1 191	821	11,7%	66,4%	138,79	Excedeu
marca	MERCEDES BENZ	2017	13 529	10 373	255,21	1 352	1 583 198	724	1 171	939	13,0%	59,8%	152,63	Excedeu
tipo veiculo	01-LIGEIRO	2017	153 809	113 096	211,52	12 015	14 175 654	504	1 180	1 132	10,6%	59,3%	125,34	Excedeu
idade condutor	21-25	2017	6 125	3 296	248,97	464	548 134	459	1 181	776	14,1%	66,8%	166,30	Excedeu
	0-2	2017	4 632	2 435	248,53	372	584 994	649	1 573	1 721	15,3%	96,7%	240,23	Excedeu
anos carta	8-10	2017	12 452	8 411	194,59	958	1 143 499	476	1 194	862	11,4%	69,9%	135,96	Excedeu
antiguidade apolice	3	2017	30 073	25 757	218,14	2 735	3 255 786	689	1 191	954	10,6%	57,9%	126,41	Excedeu
antiguidade apolice	> 10	2017	95 458	46 043	181,08	5 170	6 247 871	250	1 208	933	11,2%	74,9%	135,70	Excedeu
idade veiculo	0-2	2017	15 442	10 578	315,48	1 488	2 412 804	887	1 621	2 639	14,1%	72,3%	228,10	Excedeu
idade veiculo	3-4	2017	11 790	8 797	321,17	1 352	1 936 489	961	1 432	1 376	15,4%	68,5%	220,14	Excedeu
idade veiculo	5-6	2017	14 682	11 047	266,30	1 347	1 938 287	770	1 438	1 015	12,2%	65,9%	175,45	Excedeu

Com a análise da tabela 5.4. o gestor consegue rapidamente saber quais os segmentos onde tem de agir o quanto antes, de forma a melhorar o resultado da sua carteira. No caso em estudo, esses segmentos foram a zona geográfica Grandes Cidades, Litoral e Ilhas, os distritos Lisboa e Porto, as marcas Audi, BMW, Citroen e Mercedes Benz, os veículos ligeiros, os condutores com idades entre os 21 e os 25 anos de idade, os condutores com carta até 2 anos e entre os 8 e 10 anos de carta, apólices que estão há 3 anos em carteira e apólices com mais de 10 anos e ainda veículos com até 6 anos, uma vez que são estes segmentos que excedem os limites definidos pelo gestor para o rácio de sinistralidade, frequência e custo médio, em simultâneo.

## 5.2. Programa para criar grupos

De forma a facilitar a identificação de níveis de variáveis com comportamento idêntico e a diminuir o número de níveis, foi construído um programa em Excel que cria grupos, ou seja, agrupa os níveis da variável que se comportam de forma estatisticamente semelhante.

A base de cálculo deste programa é um modelo de análise de variância, sendo que, numa fase inicial, é calculada uma tabela ANOVA, indicando o valor da estatística de teste e o respetivo valor-P. Quando o resultado da análise da variância indica que existem diferenças entre os valores médios dos diferentes níveis da variável, o programa permite a criação de grupos homogéneos através do teste sobre a hipótese linear.

Para a criação dos grupos é construída a matriz de restrições, C, de modo a que as médias das populações pertencentes a um mesmo grupo sejam iguais. Assim, por exemplo, se quisermos testar se as médias se dividem em 2 grupos, em que o grupo 1 inclui as populações 1,2,3 e 4 e o grupo 2 tem as populações 5,6 e 7, a hipótese a testar é dada por:

$$H_0: \mu_1 = \mu_2; \mu_2 = \mu_3; \mu_3 = \mu_4; \mu_5 = \mu_6; \mu_6 = \mu_7.$$

Neste caso a matriz de restrições  $C$  que corresponde a esta hipótese é dada por:

$$C = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \quad (5.1)$$

Em consequência, a estatística de teste é dada por:

$$F = \frac{(C\hat{\mu})^T [C(X^T X)^{-1} C^T]^{-1} (C\hat{\mu}) / q}{SQ_e(C) / (n - p)}$$

e segue uma distribuição  $F$  com  $q$  graus de liberdade no numerador e  $n-p$  no denominador. Pode-se provar que esta formulação da estatística de teste é igual à que foi anteriormente apresentada, baseada na diferença das somas de quadrados de resíduos entre o modelo reduzido e o modelo completo, e é a mais adequada ao teste relativo à divisão das populações em grupos com médias iguais.

O programa foi construído de forma a ser o mais intuitivo possível para o utilizador final. O utilizador final apenas precisa de introduzir, numa fase inicial, uma tabela com estrutura semelhante às que foram apresentadas inicialmente no subcapítulo 4.3, o nome da variável em estudo, o período em estudo e ordenar os custos médios do menor para o maior. Depois recebe uma informação sobre a igualdade dos grupos. Caso haja diferenças entre os grupos, o utilizador precisa apenas de ir tentando dividir os níveis da variável pelos diferentes grupos, associando a cada variável o grupo em que considera ser o correto até o programa informar que chegou a um número de grupos adequado. Assim no final é apresentado ao utilizador a estatística de teste  $F$ , o valor  $p$  e a conclusão sobre se o número de grupos é adequado.

Depois de encontrado o número de grupos adequado, o programa devolve, para cada grupo, o respetivo custo médio.

Para exemplificar os passos que têm de ser dados pelo utilizador, foi preenchida a tabela com a variável distrito e para o ano de 2017. Na figura 5.1 é apresentada a tabela inicial que o utilizador tem de preencher para o caso do exemplo já referido.

Ano		2017		Mês		12				
Distrito	Nº Apóli	Exposição	Prémio Mé	Nº Sinistr	Custo Sinistr	Custo Médio	Desvio padrão do cus	Frequência	Rácio Sinistralda	Prémio Risc
Portalegre	5 128	3 823	160,9	332	303 477	914	489	8,69%	49,33%	79,4
Bragança	3 737	2 724	146,8	203	188 132	924	458	7,47%	47,06%	69,1
Guarda	4 079	2 925	155,8	220	203 805	926	415	7,52%	44,74%	69,7
Beja	4 021	2 835	161,5	234	226 583	967	628	8,27%	49,50%	79,9
Castelo Branco	11 120	8 016	185,7	876	856 711	978	660	10,92%	57,56%	106,9
Coimbra	14 057	10 730	189,8	1 168	1 169 416	1 001	638	10,89%	57,42%	109,0
Viana do Castelo	15 502	11 891	183,0	1 190	1 216 079	1 022	636	10,01%	55,89%	102,3
Viseu	14 663	10 872	162,5	865	902 067	1 043	764	7,95%	51,06%	83,0
Madeira	4 465	3 289	150,7	229	243 083	1 062	380	6,96%	49,03%	73,9
Setúbal	8 909	6 111	206,5	638	685 814	1 075	551	10,44%	54,36%	112,2
Aveiro	18 246	13 448	199,3	1 312	1 498 560	1 142	811	9,76%	55,92%	111,4
Braga	28 508	21 619	213,4	2 219	2 537 836	1 144	846	10,26%	55,01%	117,4
Leiria	14 687	11 278	200,3	1 189	1 363 816	1 147	785	10,54%	60,38%	120,9
Porto	51 579	38 205	223,2	4 588	5 467 192	1 192	1 529	12,01%	64,10%	143,1
Açores	1 730	1 143	145,6	66	79 062	1 198	312	5,77%	47,47%	69,1
Faro	15 017	9 968	167,8	851	1 081 233	1 270	627	8,54%	64,62%	108,5
Évora	7 550	5 360	173,0	458	590 382	1 290	2 125	8,54%	63,66%	110,2
Lisboa	18 211	13 517	265,0	1 503	2 064 087	1 374	891	11,12%	57,61%	152,7
Santarém	6 989	5 248	189,9	524	722 637	1 380	1 273	9,98%	72,51%	137,7
Vila Real	7 029	4 990	175,6	479	671 804	1 404	1 718	9,59%	76,68%	134,6

Figura 5.1: Tabela inicial preenchida com a variável distrito

A tabela ANOVA e respetiva conclusão sobre igualdade dos grupos é apresentada na figura 5.2. As hipóteses que estão a ser testadas na tabela ANOVA apresentada na figura 5.2 são:

$H_0$ : Os distritos não apresentam diferenças entre si. vs  $H_1$ : Os distritos apresentam diferenças entre si.

**Tabela ANOVA**

Fonte de Variação	Soma de Quadrados	Graus de Liberdade	Média de Quadrados	Estat.-F
Distrito	297 343 538,4	19	15 649 659,9	13,55
Erro	22 088 956 941,0	19 123	1 155 098,9	<b>Valor-P</b> 2,052E-43
Total	22 386 300 479,4	19 142	1 169 486,0	

**Conclusão:** As/Os diferentes Distrito apresentam diferenças entre si

Figura 5.2: Tabela ANOVA e Conclusão

Como podemos ver pela conclusão apresentada na figura 5.2, os distritos apresentam diferenças entre si, uma vez que o valor p é inferior a todos os níveis de significância usuais, pelo que a hipótese  $H_0$  é rejeitada. Assim, é necessário passar ao próximo passo para descobrir quais os distritos com custo médio semelhante. Na figura 5.3 é apresentada a estatística F, o valor p e a conclusão para o número de grupos apresentados na figura 5.5.

Estatística-F: 1,041  
 Valor-P: 0,408  
 Conclusão: Encontrou o número de grupos ideal

Figura 5.3: Conclusão sobre o número de grupos encontrado

Caso o número de grupos não seja adequado a mensagem apresentada é a que podemos ver na figura 5.4.

Estatística-F: 1,648  
 Valor-P: 0,049  
 Conclusão: O número de grupos encontrado não é suficiente

Figura 5.4: Conclusão quando o número de grupos não é adequado

As conclusões apresentadas nas figuras 5.3 e 5.4 são uma simples interpretação do respetivo valor p, uma vez que o teste F apresentado corresponde a testar se os grupos são semelhantes entre si, pelo que se o valor-p for superior ao nível de significância, considerado como sendo 0.1, não rejeitamos a hipótese de os grupos serem semelhantes entre si, logo, obtivemos um número de grupos adequado. Caso contrário, ou seja, caso o valor-p seja superior ao nível de significância, rejeitamos a hipótese de os grupos serem semelhantes entre si, pelo que o número de grupos não é suficiente.

# CUSTOS

Ano / Ano e Mês

12/2017

Distrito	Grupo	Custos	Médias	Desv.-pad.	Variância	Nºsinistros
Portalegre	1	303 477	914	489	238 933	332
Bragança	1	188 132	924	458	210 010	203
Guarda	1	203 805	926	415	171 949	220
Beja	1	226 583	967	628	394 819	234
Castelo Branco	1	856 711	978	660	435 214	876
Coimbra	2	1 169 416	1 001	638	407 125	1 168
Viana do Castelo	2	1 216 079	1 022	636	404 405	1 190
Viseu	2	902 067	1 043	764	583 597	865
Madeira	2	243 083	1 062	380	144 600	229
Setúbal	2	685 814	1 075	551	303 753	638
Aveiro	3	1 498 560	1 142	811	657 000	1 312
Braga	3	2 537 836	1 144	846	715 943	2 219
Leiria	3	1 363 816	1 147	785	615 828	1 189
Porto	3	5 467 192	1 192	1 529	2 337 006	4 588
Açores	3	79 062	1 198	312	97 296	66
Faro	4	1 081 233	1 270	627	393 608	851
Évora	4	590 382	1 290	2 125	4 514 787	458
Lisboa	4	2 064 087	1 374	891	794 230	1 503
Santarém	4	722 637	1 380	1 273	1 621 770	524
Vila Real	4	671 804	1 404	1 718	2 950 135	479
<b>Total</b>		<b>22 071 777</b>	<b>1 153</b>			<b>19 143</b>

Figura 5.5: *Output* para a construção de grupos

A tabela 5.5 apresenta a tabela final onde se encontram o custo médio de cada grupo, para os grupos apresentados na figura 5.5.

Tabela 5.5: Tabela Final

Grupo	Custo Médio
1	953,42
2	1 030,97
3	1 167,72
4	1 345,19

## 6. Sistemas Alternativos de Controlo de Risco

Neste capítulo o objetivo foi encontrar medidas alternativas que nos permitissem controlar o risco em carteira, de forma a melhorar o rácio de sinistralidade, através da redução de custos e do aumento de prémios, o que irá consequentemente melhorar o resultado da carteira.

Iniciou-se o estudo em duas temáticas: (1) impacto dos *Advanced Driver Assistance Systems* (ADAS) no seguro automóvel noutros países, uma vez que ainda não existem estudos para Portugal e (2) Telemáticas. Tanto os ADAS como as Telemáticas poderão permitir às Companhias apresentar prémios mais ajustados e melhorar resultados.

### 6.1. ADAS

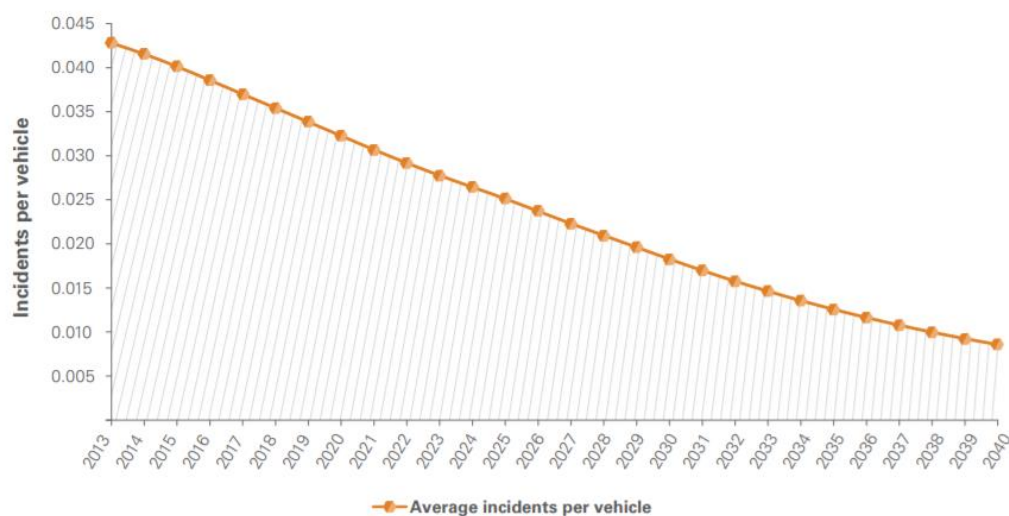
Os ADAS, que em português pode ser traduzido como Sistemas Avançados de Apoio à Condução, são sistemas desenhados para ajudar os condutores a realizar as mais diversas manobras. Estes sistemas podem enviar sinais sonoros e/ou luminosos ao condutor para o alertar sobre determinada situação, sendo assim o condutor informado da necessidade ou urgência de realizar determinada ação. Podem também os próprios sistemas agir, antecipando ou substituindo a ação do condutor.

Existem diversos sistemas avançados de apoio à condução, tais como os sensores de estacionamento, a câmara de marcha atrás, o alerta de colisão, a travagem automática, o detetor de ângulo morto, o aviso de desvio da trajetória e o sistema de luzes inteligente.

Os sensores de estacionamento ajudam o condutor a estacionar, emitindo sinais sonoros com diferentes intensidades consoante a distância da parte da frente e/ou da traseira do veículo a outro veículo ou a outro objeto. O alerta de colisão emite um sinal sonoro e luminoso quando o condutor se aproxima demasiado do veículo à sua frente. A travagem automática vem geralmente associado ao alerta de colisão, em que caso o condutor não apresente reação ao aviso de colisão o veículo trava automaticamente. O detetor de ângulo morto exibe um alerta visual nos espelhos retrovisores exteriores caso seja detetada a presença de um veículo e se o condutor acionar o sinal de mudança de direção é emitido um sinal sonoro. O aviso de desvio da trajetória emite um sinal ao condutor e recua o volante lentamente caso o veículo se desvie da faixa de rodagem.

Através do estudo realizado pela KPMG (Klynveld Peat Marwick Goerdeler) nos Estados Unidos intitulado “*Marketplace of Change: Automobile Insurance in an Era of Autonomous Vehicles*,” podemos concluir que estes sistemas diminuem a frequência, tal como é apresentado no gráfico 6.1, mas têm custos médios mais elevados, tal como apresentado no gráfico 6.2, o que é expetável uma vez que estes sistemas utilizam sensores, o que torna a reparação mais cara.

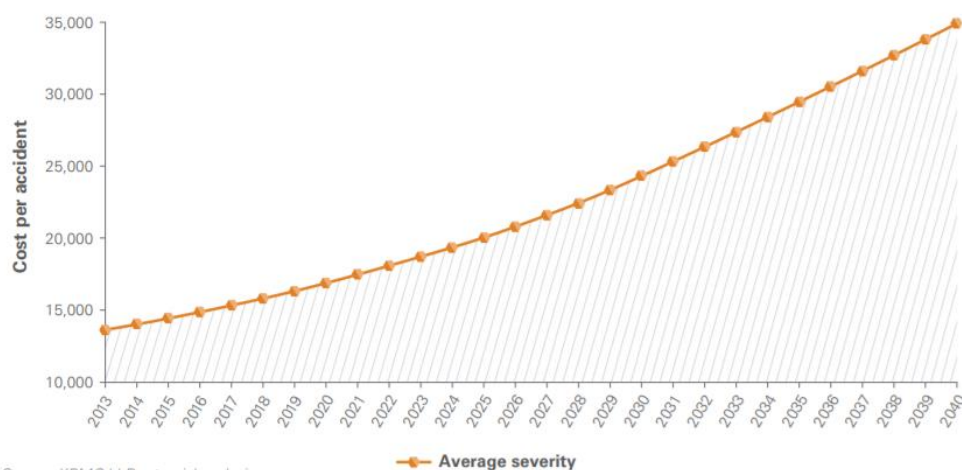
Accident frequency per vehicle by year (baseline scenario)



Source: KPMG LLP actuarial analysis

Gráfico 6.1: Evolução da frequência (ADAS)

Severity per accident



Source: KPMG LLP actuarial analysis

Gráfico 6.2: Evolução do custo médio (ADAS)

Usando os valores apresentados nos gráficos 6.1 e 6.2, podemos calcular o prémio de risco, de forma a saber se estes sistemas são ou não uma boa aposta para a Companhia e se assim pode a Companhia ajustar os prémios. Pelo que para 2018 temos  $16.000 \times 0,035 = 560$ , para 2025 o prémio de risco já será  $20.000 \times 0,026 = 520$  e no fim da previsão efetuada pela KPMG, ou seja, em 2040 será  $35.000 \times 0,0075 = 263$ . Concluimos assim que estes sistemas irão fazer com que o prémio de risco vá diminuindo ao longo dos anos, como consequência da redução da frequência e do aumento do custo médio.

O estudo efetuado pelo *Highway Loss Data Institute* nos Estados Unidos intitulado “*General Motors collision avoidance features*” realizado para alguns dos sistemas (alerta de colisão, alerta de colisão e travagem automática, sensores de estacionamento, câmara de marcha-atrás e sensores de estacionamento e câmara de marcha-atrás em conjunto), permite-nos saber o impacto na frequência dividido por cobertura, RC (Responsabilidade Civil) e colisão, tal como é apresentado na tabela 6.1.

Tabela 6.1: Impacto na frequência de sinistralidade

Cobertura	Alerta de Colisão	Alerta de Colisão e Travagem Automática	Sensores de estacionamento	Câmara de marcha atrás	Sensores de estacionamento e câmara de marcha atrás
<b>Colisão</b>	-4,1%	-5,4%	-4,3%	-2,0%	-7,1%
<b>RC</b>	-1,7%	-9,2%	-7,5%	0,4%	-16,6%

Podemos assim concluir que os sistemas quando combinados apresentam um impacto superior, como podemos ver quando juntamos os sensores de estacionamento com a câmara de marcha-atrás em que o impacto na frequência é de -16,6% em RC e -7,1% em colisão. O mesmo acontece quando juntamos o alerta de colisão com a travagem automática em que temos um impacto na frequência de -9,2% em RC e -5,4% em colisão. Conclui-se também que o impacto é maior em RC, ou seja, é maior contra terceiros.

## 6.2. Telemáticas

A telemática é a tecnologia de envio, receção e armazenamento de informações relacionadas com objetos remotos, neste caso veículos.

O seguro automóvel com recurso a estas tecnologias utiliza um dispositivo telemático que é instalado no veículo. Esse aparelho utiliza tecnologias de GPS e GSM/GPRS e um acelerómetro de alta precisão e memória para recolha e registo de dados relevantes. Contém também um cartão SIM que é utilizado para transferir os dados recolhidos para a seguradora.

Existem dois tipos de seguro automóvel possíveis com as telemáticas, que são *Pay As You Drive* e *Pay How You Drive*.

No seguro *Pay As You Drive*, o segurado paga em função dos quilómetros que faz por ano e no seguro *Pay How You Drive* o segurado paga em função do seu tipo de condução, por exemplo com uma condução mais agressiva (ex: travagens mais bruscas) pagam mais de prémio.

As telemáticas trazem vantagens para as seguradoras, tais como:

- **Melhor seleção de risco**, uma vez que este sistema é pouco apelativo para maus riscos e ajuda a reduzir situações de fraude.
- **Seguros personalizados**, pois os prémios são definidos em função do número de quilómetros ou em função do tipo de risco observado.
- **Serviço de valor acrescentado**, já que este sistema permite à seguradora ter mais informações sobre o condutor, o que lhe permite conhecer melhor os seus clientes e as suas necessidades. Poderá, assim, vender outros seguros, através de vendas cruzadas.
- **Melhor controlo de sinistros**, uma vez que através do dispositivo de telemática a seguradora sabe da ocorrência do sinistro no momento de ocorrência do mesmo.

Conclui-se assim que com telemática é possível a Companhia ajustar o prémio ao cliente, de acordo com o seu comportamento. A Companhia pode também, com este sistema, ter todas as informações sobre o cliente em tempo real e conhecer todos os sinistros do cliente.



## 7. Conclusão

O estágio na Companhia, permitiu-me compreender como os conceitos adquiridos durante a parte curricular do mestrado em Matemática Aplicada à Economia e Gestão podem ser aplicados num contexto empresarial, neste caso no âmbito de uma Companhia de seguros. Este aspeto foi particularmente interessante, pois foi feita uma ligação com as disciplinas de seguros que tinha frequentado ao longo da parte curricular.

Durante a realização deste estágio e, mais especificamente, deste projeto foi sendo perceptível que os conceitos estatísticos estão sempre presentes no dia-a-dia de qualquer departamento empresarial que trabalhe com análise de dados.

Com o aumento da competitividade do mercado, torna-se cada vez mais importante que as Companhias de seguros acompanhem a sua carteira, de forma a tomarem as melhores decisões do ponto de vista estratégico e comercial. Nesta perspetiva, é essencial terem diferentes estudos da sua carteira, pois grande parte desta será renovada de ano para ano.

Sabe-se que, com a melhoria socioeconómica do país, cada vez mais carros andam em circulação e que os custos médios de sinistros têm vindo a aumentar, pelo que é muito importante conseguir-se manter a frequência e o custo médio estável de forma manter a carteira equilibrada. Por outro lado, um aumento de prémios, também permite obter melhores resultados no final do ano. Pelo que a combinação de frequência, custo médio e prémio anual são fatores muito importantes numa seguradora.

Uma análise cuidada ao produto ganha importância quando a Companhia de seguros pretende escolher o risco que quer comercializar e manter em carteira. Aqui vê-se que tipos de variáveis independentes produzem melhores resultados para a Companhia. São essas características de risco que se pretende ter em carteira. Contudo não basta vender, mas sim vender ao preço ajustado. Numa fase posterior, são necessários métodos que permitam rapidamente identificar quais os segmentos onde agir.

Tal como em todas as áreas, o grande objetivo da Companhia é obter o melhor resultado possível no fim do ano, pelo que, muitas vezes a Companhia está disposta a perder um determinado número de riscos a fim de obter um melhor resultado.

Neste trabalho final de mestrado fez-se uma análise cuidada à carteira automóvel dos últimos 4 anos de forma a ser possível conhecer o passado recente da mesma e assim ser possível encontrar métodos que permitam melhorar a carteira automóvel da Companhia. Construíram-se dois programas que ajudam no controlo de risco em carteira. O primeiro programa permite que o gestor defina os limites que considera adequados e necessários para o rácio de sinistralidade, para a frequência de sinistralidade e para o custo médio e assim consiga saber quais os segmentos onde tem de agir rapidamente. O segundo programa desenvolvido permite ao gestor encontrar grupos com igual custo médio, de forma a poder segmentar melhor a sua carteira.

Foi também interessante perceber como é que a Companhia se comporta em relação ao mercado segurador, e perceber a tendência de mercado em relação ao risco que produz bom resultado para a Companhia.

A combinação entre o resumo comparativo anterior e a análise de mercado fez-me também concluir que apesar de se tratar de uma análise ao passado não podemos descuidar o futuro e o risco que continua a entrar em carteira, pois o que antes era um bom risco, rapidamente se pode tornar num mau risco.

Assim sendo é muito importante para as Companhias de seguros terem sistemas alternativos de controlo do risco. É aqui que os ADAS e as telemáticas se tornam um elemento chave no futuro da seleção e controlo de risco. Com o avançar das tecnologias, o custo destes equipamentos tende a diminuir e os benefícios que estes sistemas trazem para a Companhia de seguros tende a ser enorme. A diminuição da frequência de sinistralidade pode fazer com que qualquer Companhia de seguros passe a obter um melhor resultado.

Mas enquanto estas tecnologias não estão muito avançadas em Portugal, cabe a cada Companhia conseguir seleccionar o risco da melhor maneira, pelo que deve ser feita uma análise cuidada aos dados com a menor periodicidade possível com o intuito de detetar novos comportamentos de risco.

## 8. Referências

### 8.1. Referências Bibliográficas

- Daniel, Luís. (2009) ABC dos Seguros
- Guedes-Vieira, Manuel (2012) Introdução aos seguros
- Gilberto, Fernando (2012) Manual Prático dos Seguros
- Pestana, Dinis D., Velosa, Luís F. (2010) Introdução à Probabilidade e Estatística
- Murteira, B., Antunes M. (2012) Probabilidades e Estatística Volume II
- Apontamentos de apoio à disciplina de Atividade Seguradora do Mestrado em Matemática Aplicada à Economia e Gestão
- Alpuím, Teresa. Apontamentos de apoio à disciplina de Probabilidade e Estatística do Mestrado em Matemática Aplicada à Economia e Gestão da FCUL.
- Alpuím, Teresa. Apontamentos de apoio à disciplina Modelos Lineares do Mestrado em Matemática Aplicada à Economia e Gestão
- KPMG. (2015) Marketplace of Change: Automobile Insurance in an Era of Autonomous Vehicles,
- Highway Loss Data Institute. (2017) General Motors collision avoidance features.
- Swiss Re (2017) Unveiling the full potential of telematics How connected insurance brings value to insurers and consumers: An Italian case study.
- OkTeleseguros. Guia da Assistência GPS

### 8.2. Referências Sitográficas

- Associação Portuguesa de Seguradores. Disponível em: <https://www.apseguradores.pt> [Acesso em 2018/05]
- Autoridade de Supervisão de Seguros e Fundos de Pensões. Disponível em: <https://www.asf.com.pt> [Acesso em 2018/05]
- Sistemas Avançados de Apoio à Condução, IMTT. Disponível em: [http://www.imt-ip.pt/sites/IMTT/Portugues/EnsinoConducao/ManuaisEnsinoConducao/Documents/Fichas/FT\\_SistemasApoioConducao.pdf](http://www.imt-ip.pt/sites/IMTT/Portugues/EnsinoConducao/ManuaisEnsinoConducao/Documents/Fichas/FT_SistemasApoioConducao.pdf) [Acesso em 2018/09]
- Sistemas de Assistência à Condução, Sapo. Disponível em: [https://www.sapo.pt/noticias/motores/sistemas-de-assistencia-a-conducao-conheca-os\\_590cb53891857c7202d30ed6](https://www.sapo.pt/noticias/motores/sistemas-de-assistencia-a-conducao-conheca-os_590cb53891857c7202d30ed6) [Acesso em 2018/09].
- ADAS. Disponível em: <https://www.lifewire.com/advanced-driver-assistance-systems-534859> [Acesso em 2018/09]